Understanding the crystal landscape of primary sulfinamides and predicting cocrystallization

Lorraine M. Crowley, B.Sc

A thesis presented for the degree of
Doctor of Philosophy
to
THE NATIONAL UNIVERSITY OF IRELAND, CORK

Department of Chemistry,
University College Cork

Supervisor: Dr Simon Lawrence
Head of Department: Dr Justin Holmes

April 2017

I hereby confirm that the body of work described within this book, for the degree of Doctor of Philosophy, is the result of research carried out by me, within University College Cork, under the supervision of Dr Simon Lawrence, and that this thesis has not been previously reproduced, in whole or in part, at University College Cork, or elsewhere, to the best knowledge of the author, prior to the presentation of this degree.

Signed:_____

Lorraine Mary Crowley

Date:_____

Acknowledgements

The completion of this thesis would have been impossible without the help and dedication of many people. Firstly, I would like to express my gratitude to my supervisor Dr Simon Lawrence and also to Prof. Anita Maguire for the guidance and encouragement that I received throughout the completion of this thesis, particularly in the final year, as many of your evenings were sacrificed for corrections. Thank you for taking me on as a student and for helping me through the challenges, I am forever grateful.

I would like to sincerely thank the technical staff in UCC for their help over the years. Thanks to Dr Dan McCarthy and Dr Lorraine Bateman for their help with NMR; Dr Florence McCarthy and Mick O'Shea for mass spec; Barry O'Mahony for elemental analysis; Derry Kearney for fixing all of the glassware that I broke; Donnacha O'Connell for his help with demonstrating; and to Tina Kent, Christine Dennehy and Denis Duggan for their services and help (and all the great chats); to Noel Browne for solvent delivery and to Dr Matthias Jauch for IT services (and for trying to resurrect my laptop). Thanks to the Irish Research Council for funding my PhD research.

The memories of my days in UCC are filled with great friends who made the days so fun and enjoyable. I would like to thank my lab mates Carla, Abhi, Uday, Balla who made the days in the lab brighter and especially to Stephen, who always encouraged me to keep going (and who's singing made the days in the lab more entertaining, especially at Christmas). Thanks so much to Aaran and Aoife for being the best two desk-neighbours I could ever ask for (you'll love the Mac eventually, Aaran), and to all of the '10am tea group' for the great conversations. Thanks to Dr Declan Gavin for always making time for my questions and for being a genuinely great friend. I won't ever be able to express enough thanks for the help that I received from Dr Denis Lynch over the years, you were never too busy to help (and I'll never throw out the mother liquor!).

To Claire, for being with me through every second of this, you are the definition of a friend, never too busy for an overly-foamy soy latte and a chat, we did this together and I'll forever be in your debt.

A special thanks goes to my colleagues in Janssen, who have proved that all of this effort was so worth it, to John and Michelina for giving me a great job; and to all my friends in the business units who fill the days with so much craic (and a small bit of work!), it's a privilege to work with you all. A special thanks goes to Ray, for proofreading this thesis (and the laughs over posh coffee).

To Mam and Dad, to whom I dedicate this thesis, for encouraging my education from day one, and for pushing me through the journey - always so confident that I could finish it (although there were days you were sure ye had Sheridan on your hands!), ye were never too busy to help me, and always there to talk, thank you both so much. To Stephen, for the weekly phone calls and for being so proud of his big sister, and to Patrice for being such a good little sister (and for cleaning my apartment when I was so busy).

And finally, Dan, there are no words to express my gratitude to you for sticking with me through this. You listened to every single complaint, consoled every failure, and celebrated even the smallest of wins with me. You made me believe I could do it, and it just wouldn't have happened without you, thank you *so* much.

Abstract

This thesis focuses on design principles in crystal engineering, from the development of our knowledge of understudied functional groups, to the synthesis of cocrystals.

Chapter one is a brief introduction to the concepts and development of the field of crystal engineering over the past 30 years, including the definition and utilisation of the hydrogen bond in the synthesis of multicomponent materials, and reviewing the cocrystallization literature involving sulfur functional groups.

Chapter two describes the various synthetic strategies utilised in the synthesis of a library of primary aryl sulfinamides, an under-studied sulfur functional group from a solid state chemistry perspective. Their solid-state properties are examined, with focus on the robust supramolecular synthons that they exhibit. Finally, this chapter describes their hydrolysis, examining the relationship between their hydrolytic sensitivity and solid state structure.

Chapter three outlines the preparation of novel solid forms of the active pharmaceutical material Salsalate. These include the newly observed phenomenon of reactive cocrystallization.

Chapter four outlines the development of an experimental data matrix and machine learning algorithm for the computational prediction of cocrystal formation, and the use of this software to rank and compare the probable cocrystals from an external test set. This work was conducted in collaboration with Prof. Richard Cooper and Jerome Wicker of the University of Oxford, UK. The solid-state properties of many of the novel cocrystals identified in the development of the experimental data matrix are described, with focus on the presence of recurring and predictable hydrogen bonding motifs observed in these systems.

Chapter five examines the future work that can be envisaged building on the work described in the earlier chapters.

List of Abbreviations

Abbreviation	Interpretation
ABq	AB quartet
acac	acetylacetonate
API	Active Pharmaceutical Ingredient
Ar	Aromatic
ATR	Attenuated Total Reflectance
br s	broad singlet
Bz	Benzyl
CCDC	Cambridge Crystallographic Data Centre
$CDCl_3$	Deuterated Chloroform
CH_2Cl_2	Dichloromethane
$CHCl_3$	Chloroform
Cq	quaternary carbon
CSD	Cambridge Structural Database
CSD Refcodes	CSD refcodes follow the general format AAAAAA and are used to refer to crystal structures found in the CSD, suffixes of 01/02 etc. relate to polymorphs of these materials
d	doublet
DBSO	Dibenzylsulfoxide
DCC	Dicyclohexylcarbodiimide
dd	doublet of doublets
ddd	doublet of doublet of doublets
DMSO	Dimethylsulfoxide
DPSO	Diphenylsulfoxide
DSC	Differential Scanning Calorimetry
dt	doublet of triplets
e/n	electronegativity value
ee	Enantiomeric excess

Eq	Equivalents
ESI	Electron Spray Ionisation
EtOAc	Ethyl Acetate
FDA	Food and Drug Administration
GooF	Goodness of Fit
HBP	Hydrogen Bond Pairing
HIV	Human Immunodeficiency Virus
HRMS	High Resolution Mass Spectrometry
Hz	Hertz
I	Intensity
IPA	Isopropanol
IR	Infrared Spectroscopy
IUPAC	International Union of Pure and Applied Chemistry
J	coupling constant (Hz)
L	Litres
LiHMDS	Lithium Bis(trimethylsilyl)amine
Lit.	Literature result
M	Molarity (Mol L^{-1})
m	multiplet
m/z	Mass to charge ratio
MeOH	Methanol
MeOD	Deuterated Methanol
MHz	Mega Hertz
min	Minutes
mL	Millilitres
mol	moles
mp	Melting point
Mr	Molecular mass
NBS	*N*-Bromosuccinimide
NH_4Cl	Ammonium Chloride
NEt_3	Triethylamine
NMR	Nuclear Magnetic Resonance spectroscopy
obs.	Observed

***p*-**	Para
Ph	Phenyl
ppm	parts per million
PXRD	Powder X-ray diffraction
q	quartet
RT	Room Temperature
s	singlet
Sat.	Saturated
SCXRD	Single crystal X-ray Diffraction
t	triplet
td	triplet of doublets
THF	Tetrahydrofuran
TLC	Thin Layer Chromatography
SVM	Support Vector Machine
SSIP	Surface Site Interaction Point
EF	Enrichment Factor
ROC	Receiver Operating Curve
AUC	Area Under Curve
NSAID	Non-Steroidal Anti-Inflammatory Drug
GI	Gastro-Intestinal
GRAS	Generally Regarded As Safe
HSA	Hirshfeld Surface Analysis

Table of Contents

To Mam and Dad

Chapter 1

Introduction

1.1 Introduction to Crystal Engineering

1.1.1 Crystal Engineering – An Overview

Crystal engineering describes the efforts to produce solids with predefined and predictable physical properties, arising from the design of specified intermolecular interaction motifs related to the molecular structure.[1,2] Crystal engineering is a field of research that has been steadily growing since the 1950's, but has come to the forefront of academic research over the past two decades.[3,4] The term 'crystal engineering' was formally introduced to the literature by Schmidt in 1971,[5] in his report of the photodimerisation of solid materials, when he referred to 'crystal engineering' as the development of rules to govern and design molecular packing.

Although the link between physical properties and molecular packing had been made decades earlier,[6] the need for more powerful analytical technology in areas such as X-ray diffraction restricted progress in this area for a long period of time. It wasn't until the 1970's and 1980's that the field began to gain the interest of the broader academic community.[1] While the accolade for coining the term 'crystal engineering' goes to Schmidt, there is little argument among academics that some of the greatest contributions to the field of crystal engineering over the past 30 years have been made by Gautam R. Desiraju, who has been cited almost 30,000 times across 425 publications on the subject of crystal engineering/crystallography; he is referred to as 'one of the founders of organic crystal engineering'.[3] In 1989, Desiraju provided his definition of crystal engineering as 'the understanding of intermolecular interactions in the context of crystal packing and the utilisation of such understanding in design of new solids with desired physical and chemical properties'.[7] This definition has received no modification to date, and continues to be the chosen definition used by chemists and crystallographers alike.

This chemistry, taken 'beyond the molecule'[8] is termed as supramolecular chemistry (as defined by Jean Marie Lehn in 1969).[9] Desiraju likened the process of molecular design in organic chemistry to crystal engineering efforts when he

spoke of the work of Lehn in supramolecular chemistry,[10] he said 'If molecules are built by connecting atoms with covalent bonds, then solid-state supermolecules (crystals) are built by connecting molecules via intermolecular interactions' (Figure 1).

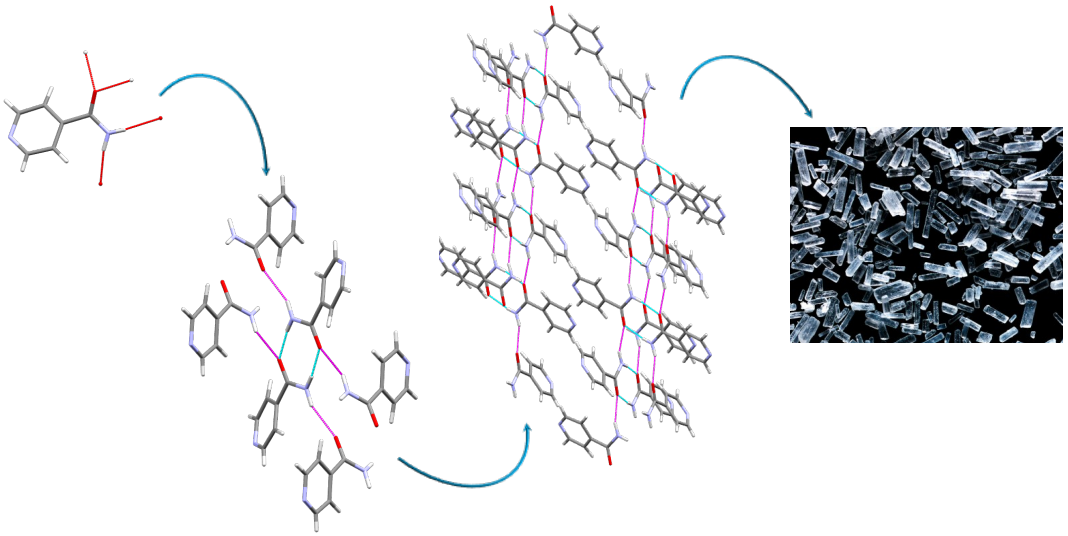


Figure 1: Supramolecular assembly via hydrogen bonds in nicotinamide [1] (CSD Refcode: EHOWIH01).[11]

Crystal engineering provides a chemist with a new avenue for control over the properties of a material. Decades of intense research in organic chemistry allows a chemist to design and synthesise target molecules, using reliable reactions such as the Suzuki and Wittig reaction to make new bonds. The field of crystal engineering has not yet progressed to providing the same level of control over the supramolecular state.[12]

The traditional organic chemistry approach is to first successfully develop an effective synthetic route to a target molecule, after which attention then goes toward the physical behavior of the synthesised material, particularly in the field of pharmaceutical materials. This has inherent issues, particularly for active pharmaceutical ingredients (APIs), the majority of which are administered as solid dosage forms, tablets in particular.[13] All too often a biochemically effective molecule is synthesised, only for its physical properties to be undesirable; poor solubility being a particularly large issue at clinical trial stage.[14,15] Knowledge of the organic solid state and crystal engineering is increasingly applied in the drug design process, thereby helping to produce effective solid materials with

desirable physical properties and preventing the above issues on an API's route to the market (Figure 2).

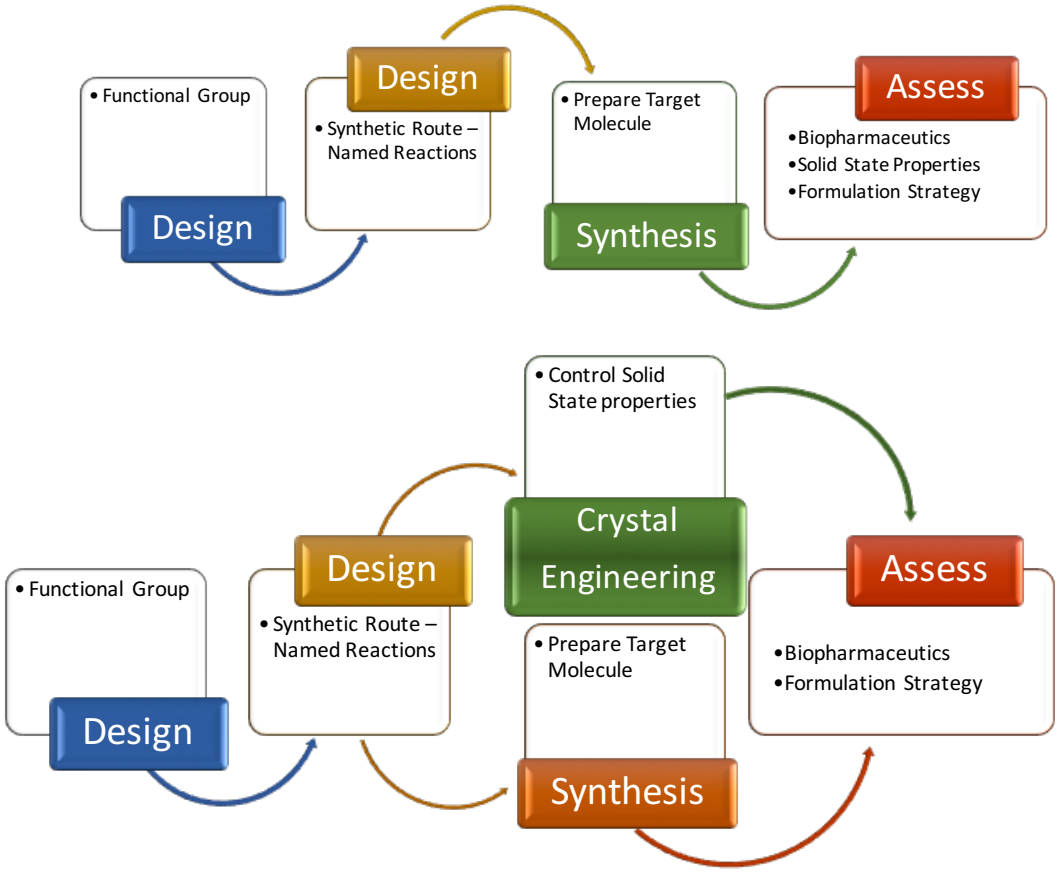


Figure 2: The normal strategy toward drug design and formulation (top) and crystal engineering incorporated in API design/formulation (bottom).

The ideal scenario would be that, given a molecular structure, the crystal structure and properties of a material would be correctly predicted.[16] John Maddox presented the challenge in 1988, when he described our inability to 'predict the structure of even the simplest crystalline solids' as a 'continuing scandal in the physical sciences'.[17] In 1994, Gavezzotti posed the question 'are crystal structures predictable?', answering with a general answer of 'no', but the qualified allowance for 'maybe'.[18] Crystal structure prediction has progressed hugely in recent times, with many different methodologies available for use,[19] for example, simulated annealing,[20] *ab initio* methods,[21] genetic algorithms,[22] and global search of the lattice energy surface.[23] Crystal structure prediction is performed in two parts, firstly, a proposed trial model of the crystal structure is generated, and this is followed by ranking of the generated structures [dispersion

corrected density functional theory (DFT-D) is a commonly used method for ranking].[24,25]

Since 1999, the Cambridge Crystallographic Data Centre (CCDC) has conducted 6 periodic blind tests of organic crystal structure prediction.[26–31] The participants are given 2-dimensional molecular structures of crystals whose structures are known but unpublished, along with the crystallization conditions, and are then required to submit their proposed crystal structures, which are then compared to the determined crystal structures for accuracy. The sixth blind test was completed in 2016, and highlighted particular success in structure prediction using Monte Carlo parallel tempering for structure generation and density functional theory (DFT) for ranking, which correctly proposed all of the five target crystal structures.

The final conclusion of this most recent test outlined concerns relating to the ever increasing computational cost of these prediction methods (up to 500,000 CPU hours in some cases) and remarked on the difficulties encountered in generating the initial model crystal structure, stating that there are still 'significant challenges' remaining for 'routine and reliable' crystal structure prediction calculations. However, the report also commented that the wide range of available methods coupled with the advancement in recent years 'bodes well' for crystal structure prediction going forward.

1.1.2 Hydrogen Bonding in the Crystal Engineering Toolbox

The major challenge in crystal engineering is to determine a reliable set of rules to predict the behaviour of a crystal from its molecular structure, thereby providing a knowledge-base for the synthesis of tailor-made solid materials, as Desiraju stated, 'the utilisation of understanding' in design.[7]

Intermolecular interactions are the building blocks in supramolecular chemistry. There are several intermolecular interactions of interest in the field of crystal engineering, including (but not limited to) van der Waals forces, π-π stacking, dipole-dipole interactions, halogen bonding and hydrogen bonding; the latter

being the interaction most commonly targeted for organic supramolecular assembly. The hydrogen bond has been described as the 'master key' to molecular recognition and 'the most important interaction type in crystal engineering'.[32,33]

The traditional understanding of a hydrogen bond is an electrostatic polar interaction, X-H···Y, between a hydrogen atom bonded to an electronegative element (X-H where X can be O, N, S or another) to another electronegative atom, Y (where Y is also O, N, S, a lone pair or another atom). A traditional hydrogen bonding interaction between water molecules is shown in Figure 3.

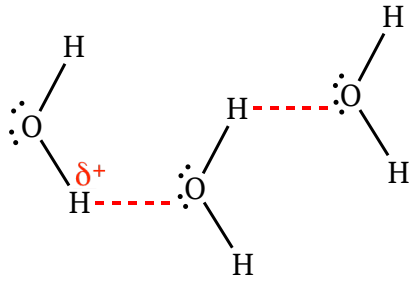


Figure 3: Traditional depiction of a hydrogen bonding interaction in water.

IUPAC have defined the hydrogen bond as 'an attractive interaction between a hydrogen atom from a molecule or a molecular fragment X–H in which X is more electronegative than H, and an atom or a group of atoms in the same or a different molecule, in which there is evidence of bond formation'.[34] This publication outlines six criteria for the formation of a hydrogen bond:

1) Formation of an electrostatic interaction typically depicted as X-H···Y-Z (where X is the donor and Y the acceptor).
2) The existence of a covalent bond between X and H.
3) Changes in the IR shift due to alteration of the X-H bond length.
4) Changes in NMR signals due to changes in the environments of X, H and Y.
5) Bond angles (typically 180°, but any angle greater than 110° is accepted).
6) Gibbs energies of the interactions (-0.2 to 40 kcal mol^{-1}).

Prior to the publication of the above definition, Jeffrey[35] published a well-accepted classification guideline for strong, moderate and weak hydrogen bonds (Table 1). Steiner referred to this table in his 2002 publication,[36] stressing that there are 'no natural borderlines in these categories' and that use of this table in

too stringent a fashion would be pointless. Steiner also referred to the use of the van der Waals cut-off as inappropriate for determination of appropriate hydrogen bond length, stating that bonds up to 3.2 Å should be included for consideration as potential hydrogen bonding. This sentiment has been echoed in the IUPAC definition,[34] and in further publications more recently.[37]

Table 1: Guidelines for hydrogen bond classification reproduced from Jeffrey.[35]

Interaction Type	Strong	Moderate	Weak
	Strongly Covalent	Mostly Electrostatic	Electrostatic /Dispersion
Bond Length [Å] H···A	1.2 – 1.5	1.5 – 2.2	> 2.2
Lengthening of X-H [Å]	0.08 – 0.25	0.02 – 0.08	< 0.02
X···A [Å]	2.2 – 2.5	2.5 – 3.2	> 3.2
Directionality	Strong	Moderate	Weak
Bond Angles [°]	170 - 180	> 130	> 90
Bond Energy [Kcal mol^{-1}]	15 - 40	4 - 15	< 4
Relative IR Shift [cm^{-1}]	25 %	10 – 25 %	< 10 %
^{1}H Downfield Shift [ppm]	14 - 22	< 14	-

An interesting statistical study was performed by Wood *et al*. in 2009,[38] where they analysed several factors relating to hydrogen bonds among the available structures in the Cambridge Structural Database (CSD).[39] One of the most compelling comparisons made here was the default settings for identification of hydrogen bonds in crystal structures for four popular software packages (Table 2).

Table 2: Default settings for popular crystallography software packages, reproduced from Wood *et al*.[38]

Software Package	H···A Distance [Å]	D-H···A Angle [°]
DIAMOND 3.1[40]	1.2 < H···A < 2.1	> 120
Mercury 2.2[41]	H···A < r_H^* + r_A^*	> 90
PLATON[42]	H···A < r_H^* + 2.0	> 110
SHELX (HTAB)[43]	H···A < r_H^* + r_A^* - 0.12	> 100

*r_H and r_A refer to the van der Waals radii of the donor and acceptor atoms (H and A).

The relatively large differences in the default settings were interpreted as indication that there is, as yet, no solid consensus on the defining parameters for a hydrogen bond, and statistical analysis of the CSD[39] was performed accordingly. In their study, the hydroxyl-pyridyl (O-H···N-C) interaction is used as a model for a 'strong' interaction, hydroxyl-carbonyl (O-H···O=C) as a moderate interaction,

and the hydroxyl-ether (O-H···O-C) as model for a 'weaker' interaction, and compared the bond lengths observed for all three. The majority of interactions were identified in the 1.7 – 2.0 Å distance range, which would be classified as moderate under Jeffrey's criteria, although the hydroxyl-pyridyl interaction is referred to as 'strong'. None of the interaction distances observed in this study would be classified as strong interactions according to Jeffrey's criteria, suggesting that the 1.5 Å tolerance is restrictive in its classification.

Interestingly, bond angle analysis revealed that 72% of identified interactions have a bond angle of 150° or greater, with just 6% at less than 120° (which would be adjusted to just 1% with a cone correction). This finding would suggest that the 110° value suggested in the IUPAC definition may be on the generous-side for realistic determination of hydrogen bonding interactions, and the team suggest that a cut-off of 135-140° may be more appropriate for intermolecular interactions. A third interesting finding of this study was the observation that interactions with bond angles of 120-140° have a substantially reduced stabilisation energy.

Despite the debate surrounding the definition(s) and cut-off points for hydrogen bonding, they are none the less the most commonly used supramolecular interaction. Around the time that Gautam Desiraju proposed his definition of crystal engineering, Margaret C. Etter proposed a set of guidelines for the strategic use of hydrogen bonding interactions in a design approach.[44] In 1991, Etter published a paper entitled 'hydrogen bonds as design elements in organic chemistry', containing a section that has come to be known as 'Etter's rules'.[45] In this publication, she built upon the work of Leiserowitz,[46] Schmidt,[47] Donahue,[48] and identified the hierarchical nature of hydrogen bonds, as well as publishing this reliable set of rules for how hydrogen bonds form in molecular solids. Etter's rules state that:

(a) All good proton donors and acceptors will be used in hydrogen bonding.

(b) Typically the best donor and best acceptor will form a hydrogen bond to each other.

(c) Where a six-membered intramolecular S(6) ring can be formed, this will form preferentially to intermolecular interactions.

Etter's rules can be observed in action in the cases of 4-aminobenzoic acid [**2**] (AMBNAC01)[49] [Figure 4] and 4-aminosalicylic acid [**3**] (AMSALA01)[50] [Figure 5].

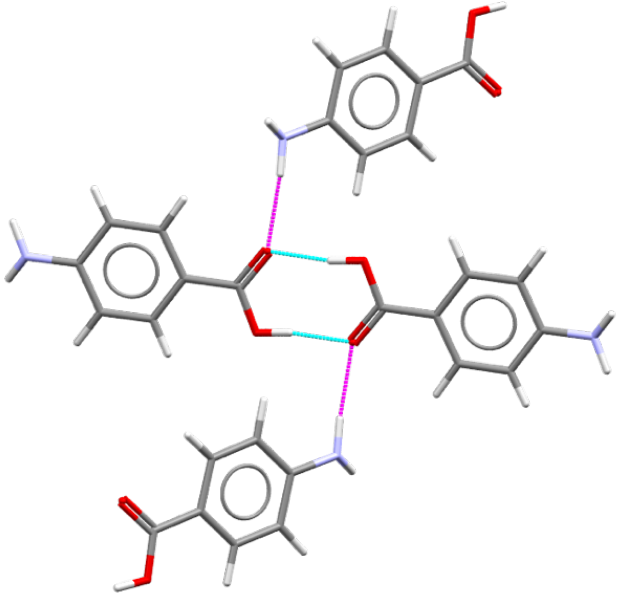


Figure 4: Hydrogen bonds observed in 2 (AMBNAC01)[49] [magenta and cyan].

AMBNAC01[49] [**2**] displays hydrogen bonding between the best donor (carboxylic acid, O-H) and the best acceptor (carboxylic acid, C=O), and also from the primary amine to the carboxylic acid in a hierarchical fashion. When the hydroxyl moiety is introduced *ortho* to the carboxylic acid in AMSALA01[50] [**3**], the intramolecular S(6) ring forms, according to Etter's rules, and a hydrogen bond forms from the primary amine to the phenolic oxygen instead, since the carboxyl oxygen is already accepting two hydrogen bonds (Figure 5).

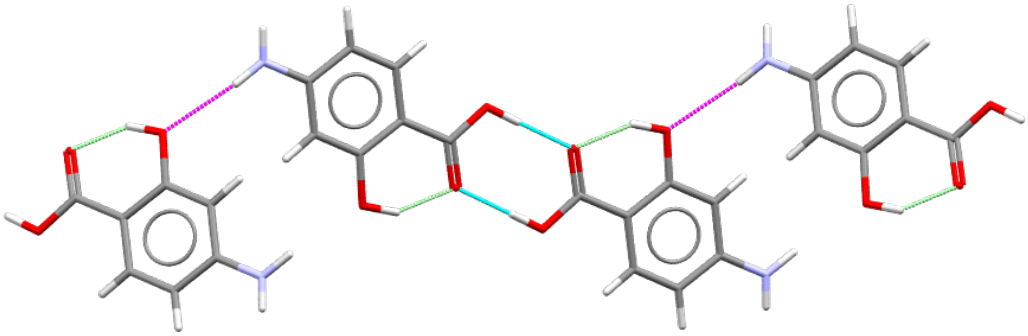


Figure 5: Etter's rules as observed in 3 (AMSALA01)[50] [hydrogen bonds in magenta, cyan and S(6) in green].

A significant piece of literature to add to the crystal engineering 'toolbox' was Hunter's publication of quantified intermolecular interactions in 2004.[51] Hunter's work published a numerical evaluation of the hydrogen bond donor and acceptor abilities of different functional groups based upon both hydrogen bond donor/acceptor constants and molecular electrostatic potential calculations, herein referred to as 'Hunter's table'. Hydrogen bond acceptor constants were based upon literature values, as derived from the work of Abraham and Platts,[52] providing an expression of the energy of a hydrogen bond interaction as the product of the positive charge on the hydrogen-bond donor (α_2^H) and the negative charge on the hydrogen-bond acceptor (β_2^H).

Hunter's table allows for a general assessment of potential pairwise hydrogen bonding interactions in the design of supramolecular architectures, cocrystals in particular, although it should be noted that the values cannot be considered *too* strictly, as 3-dimensional factors such as close packing can lead to breakdown of the hierarchichal pairwise relationship in certain cases. This is also true of Etter's rules,[45] which can relax to accommodate close packing relationships in the solid state. In Hunter's table,[51] a higher value indicates increased hydrogen bonding capacity. A subset of Hunter's table to include the functionalities relevant in this thesis is shown in Table 3.

Table 3: Hydrogen bond parameters for functional groups pertinent to this work (subset from Hunter).[51]

H-Bond Donor	α		H-Bond Acceptor	β
Aryl Bromide	1.3[b]		*Aryl Bromide*	2.2[a]
Aryl Chloride	1.3[b]		*Aryl Chloride*	2.2[a]
Aryl Fluoride	1.4[b]		*Aryl Fluoride*	1.6[a]
Pyridine	1.4[b]		*Thiol*	2.7[a]
Amine	1.5[b]		*Phenol*	2.7[a]
Thiol	1.7[b]		*Disulfide*	3.7[a]
Alcohol	2.7[b]		*Nitrile*	4.7[a]
Amide	2.9[a]		*Carboxylic Acid*	5.3[a]
Urea	3.0[a]		*Ester*	5.3[a]
Sulfinamide	3.2[b]		*Sulfonamide*	5.8[a]
Sulfonamide	2.8[b]		*Sulfone*	6.3[a]
Thioamide	3.3[b]		*Pyridine*	7.0[a]
Carboxylic Acid	3.6[a]		*Amine*	7.8[a]
a - Value based on literature values α_2^H and β_2^H			*Amide*	8.3[a]
b - Value based on molecular electrostatic potential surface.			*Sulfinamide*	8.3[a]
			Urea	8.3[a]
			Sulfoxide	8.9[a]

1.1.3 Cocrystal, Co-crystal, Molecular Complex?

A cocrystal, in the most basic sense, is a crystalline lattice that incorporates more than one discrete molecule, a multi-component crystal. The first report of a cocrystal was that of quinhydrone [**4.5**] in 1844[53] (a 1:1 cocrystal of quinone [**4**] and quinol [**5**], Figure 6), but the field has only gained the attention of the pharmaceutical industry and crucially, the FDA in very recent times. The term 'cocrystal' or 'co-crystal', as well as the appropriate definition for same, has been the subject of extensive, and ongoing, academic debate.

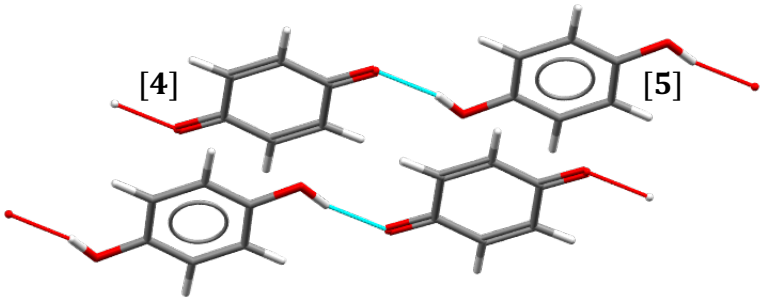


Figure 6: Crystal structure of 1:1 cocrystal of quinhydrone [4.5] (QUIDON).[54]

In 2003, Desiraju published an essay strongly arguing against the use of the term cocrystal (or co-crystal), stating that we have a 'perfectly good' term to describe multi-component crystals, *molecular complex*.[55] He also argued that the term cocrystal is 'ambiguous', and possibly 'scientifically suspect'. Shortly thereafter, Dunitz published a response in defence of the term, but insisted that the use of the hyphen (co-crystal) is 'essential'.[56]

Significant academic debate continued, with Aakeröy presenting his own personal definition in 2005,[12] including strict requirements for neutral species, reagents that are solids at room temperature, and structural homogeneity. Aakeröy's strict definition was not well received, despite his statement that the definition provided was meant for the purpose of his article only, as it excluded too many other multi-component materials. In 2011, the FDA published a draft guidance for the appropriate definition of a pharmaceutical cocrystal as '*solids that are crystalline materials composed of two or more molecules in the same crystalline lattice*'.[57] The guidance was heavily criticised for its simplicity, and for not appropriately setting cocrystals apart from salts, hydrates, or solvates.[58] The guidance was revised in 2013, describing that nonionic interactions are required for cocrystal formation and and that an API cocrystal should be considered as a 'drug product intermediate';[59] and again most recently in August 2016,[60] where cocrystals were clearly distinguished from salts, stating that the components must be neutral, likening a cocrystal to a new polymorph of an API (it cannot be considered as a completely new API).

Building upon the extensive debate provided by Aitipamula *et al*.[58] in response to the 2011 document, Grothe *et al.*[61] published a classification system to deal with the controversy. Although inelegant, the system appears to be effective in distinguishing each crystal form into three separate classes and four subclasses, without an obvious level of ambiguity (Figure 7).

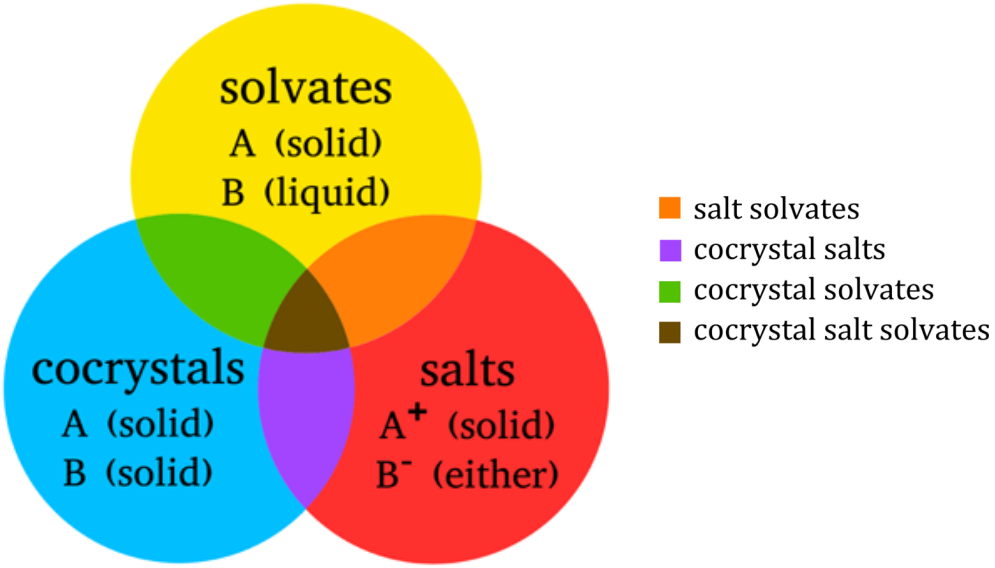


Figure 7: Classification system for multi-component solid forms adapted from Grothe *et al.*[61]

Although frustratingly complex, the level of academic (and now regulatory) debate on the subject of cocrystals should be interpreted as an indication of a mature field that is growing in importance and influence, gaining the attention of the academic heavyweights and regulatory bodies alike. The revised FDA guidance paves the way for the development of pharmaceutical cocrystals on an industrial scale, which should in turn open up a route to market for effective pharmaceutical materials that are currently unmarketable due to undesirable physical properties. The broader impact of this field is on the patients, who may gain access to new APIs and/or better drug product formulations. Cocrystal formulations have just recently made their debut to the market. In 2014, the FDA approved the first API cocrystal (dapagliflozin:propylene glycol:monohydrate, [**6**]) [Figure 8] for use in the treatment of type 2 diabetes.[62] Dapagliflozin API (market name Farxiga®) is amorphous, unstable and very hygroscopic, thus presenting significant processing issues on an industrial scale.[63] These issues are circumvented in the case of the cocrystal, while retaining the pharmaceutical activity of dapagliflozin for diabetes treatment.

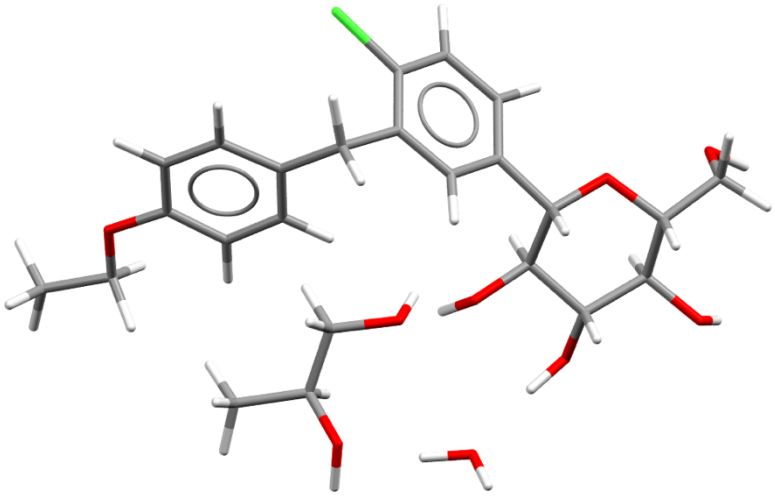


Figure 8: Structure of FDA approved API cocrystal dapagliflozin propylene glycol monohydrate [6] (CIMNUJ).[63]

1.1.4 Coformer Selection - The Use of Supramolecular Synthons

Cocrystallisation studies have successfully altered bioavailaibility and solubility,[13,64] hygroscopic stability,[65] compressibility,[66] instrinsic dissolution rate,[67] and thermal properties[68] of target materials. Crystal engineering of novel cocrystals arises from effective utilisation of our understanding of intermolecular interactions. In this context, 'supramolecular synthons' are structural motifs identified in the solid state, involving distinct, robust, and reliable intermolecular interactions, most commonly hydrogen bonds, that are repeatedly observed.[1,10] These interactions present the synthetic tools for construction of supramolecular architectures.[12] In order to design cocrystals, careful selection of a series of appropriate coformers of complimentary molecular structure to the target is essential.

Rational selection of suitable coformers involves careful assessment of the potential landscape of interactions for the target material from the literature available. These interactions (synthons) are then used as targets in the design of other novel materials.[32] Supramolecular synthons are categorized broadly into two types, homomeric (where the hydrogen bond donors/acceptors lie on the same functional group), and heteromeric (donors/acceptors on different functional groups); although homomeric can also be used to describe

interactions between molecules of the same material. The most common supramolecular synthons to be exploited in organic cocrystallization are 1) the acid-acid homodimer, 2) the amide-amide homodimer, 3) the acid-amide heterodimer, and 4) the acid-pyridyl heterodimer (Figure 9).[69]

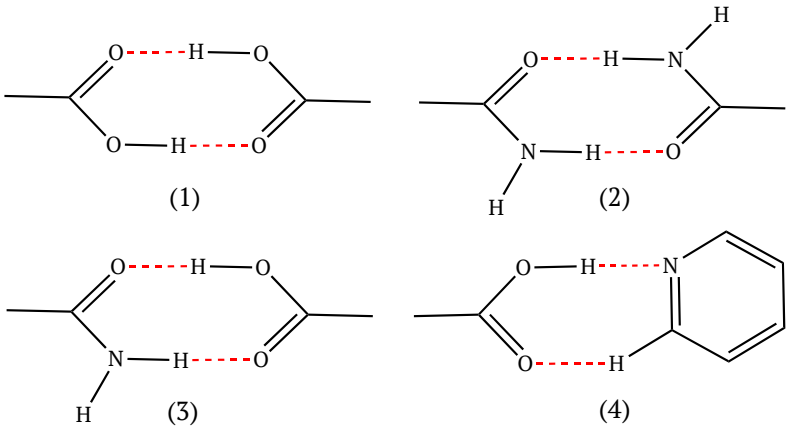


Figure 9: The most common supramolecular synthons observed in cocrystals.[69]

A concise explanation of the factors involved in knowledge-based coformer choice using hydrogen-bond pairing (HBP) was published by Pidcock and coworkers in 2014 (Figure 10).[62] Following coformer selection, preparation of cocrystals is *via* trial and error, with large sets of coformers ordinarily used to mitigate the number of failed attempts.

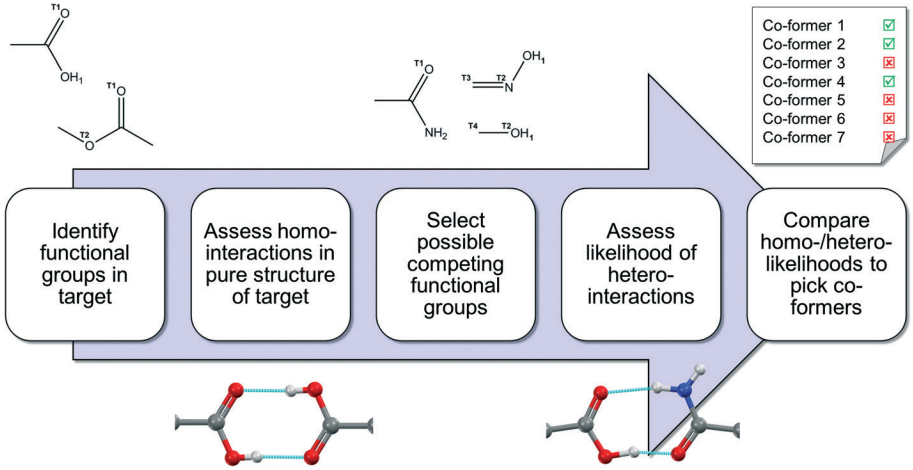


Figure 10: Rational coformer selection process as outlined by Pidcock and coworkers.[62]

The most common methodology used to generate a library of novel cocrystals is to generate a set of rational coformers, either from a list of materials commonly used within the particular research group, or a new set of materials, or from the CSD.[39] In all cases, a target suite of supramolecular synthons are kept in mind.[70]

1.1.5 Graph Set Analysis of Hydrogen Bonded Structures

Graph set analysis is a method of defining the motifs observed in hydrogen bonded systems. The modern graph set methodology has built upon the work of Wells, Hamilton and Ibers,[71,72] who used the notations *n* and *m* to define the number of bonds formed to a particular molecule and the number of molecules to which it hydrogen bonds. In 1980, Kulehsova and Zorky described the use of graph set analysis to hydrogen bonding, proposing a general symbol of $G_m^n(k)$, where *n* and *m* are as above, and *k* refers to the dimensions of the ring, if present [Figure 11]. *G* was interchanged for *I*, *C*, *L* or *F* to denote the formation of islands, chains, layers or frameworks. In these methods, the molecules were represented as vertices of a graph, with hydrogen bonds connecting the vertices.

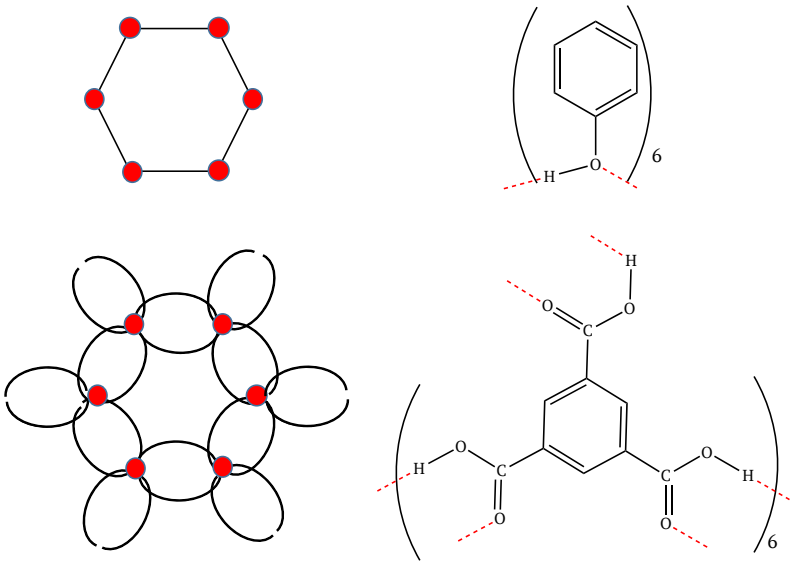


Figure 11: Graph set representation of molecules as points, as proposed by Kuleshova and Zorky,[73] reproduced from Etter.[74]

Margaret Etter *et al.*[75] proposed the use of $pG_d^a(r)$, where *p* refers to the total number of donors and acceptors in the molecule, *a* and *d* are the participating donors and acceptors, and *r* representing the ring size in the same way as *k*, above. In this case *G* could be changed for one of three specified hydrogen bonding modes, *I* (intramolecular), *C* (chain) or *R* (ring). This method was applied to aggregate structures of carboxylic acids and amides (Figure 12).

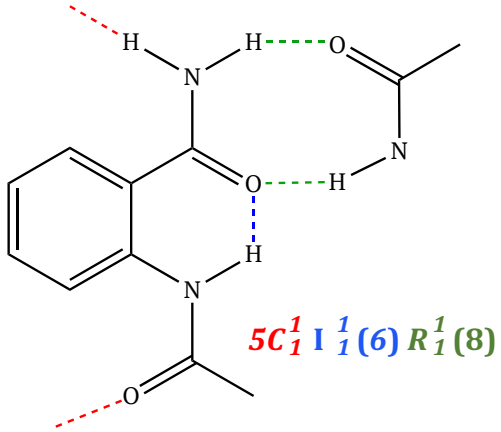


Figure 12: Graph set notation for amide hydrogen bonding patterns as proposed by Etter *et al.* (1985).[75]

Etter, Bernstein and coworkers published a number of subsequent publications on this subject.[45,76,77] The original notation was revised, suggesting the use of *C*, *R*, *S*, and *D* to describe chains, rings, intramolecular bonds, and finite bonds in place of *G*. The final graph set descriptor described by Bernstein *et al.* in 1995[78] was $G_d^a(n)$, where *n* is the number of atoms participant in the motif (Figure 13). This system allowed for effective characterization of simple unitary graph sets, where only one type of hydrogen bonding involved, as well as more complex cases in which multiple hydrogen bonds combine. Two different hydrogen bonds is referred to as a binary motif (or second level graph set), with three hydrogen bonds a ternary motif etc. This notation is currently used, and will be used in this thesis.

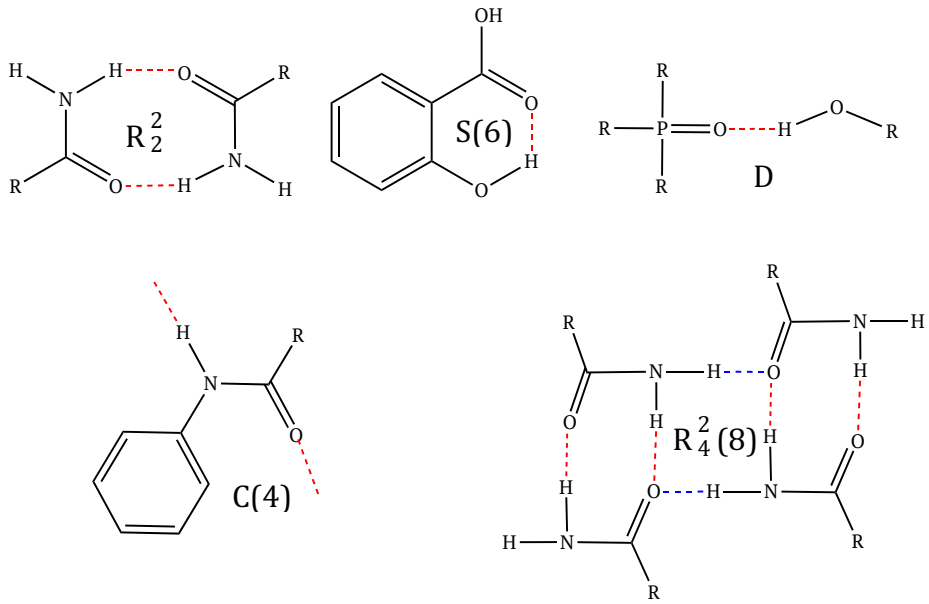


Figure 13: Finalised graph set notation for common unitary (red) and binary (red and blue, bottom right) hydrogen bonding motifs as proposed by Bernstein *et al.*[78]

The 2016 update to Mercury structure visualization software[41] includes motif search capability, which allows the user to search for simple unitary hydrogen bonding motifs, such as C(4) chains. The reported results include a representation of percentage frequency, which is the instances for which the particular motif is observed as percentage of the molecules in which that motif could form [Equation 1].

$$\%\ Frequency = \frac{No.\ of\ motif\ observed}{No.\ of\ molecules\ in\ which\ motif\ could\ form}$$

Equation 1: Calculation of % frequency as performed by Mercury.

1.2 Cocrystallization Involving Sulfur Functional Groups

Sulfur functionalitites represent an important class of organic materials, from biologically essential amino-acid residues such as cysteine and glutathione, to organic materials such as sulfonamides, sulfides, sulfoxides and sulfones. Sulfur is an important heteroatom in biologically relevant materials, incorporated into 362 FDA-approved drugs.[79] Sulfur functionalities form part of the treatments for a broad and varied series of ailments, from dry mouth (Tiopropen, a thione) to narcolepsy (Modafinil, a sulfoxide) and anti-retroviral therapy for HIV (Prezista, a sulfonamide). The main classes of sulfur functional groups including some interesting examples from organic and medicinal chemistry are illustrated in Figure 14.

With so many notable examples of sulfur functional groups, it comes as no surprise that they have received attention from a crystal engineering perspective, including a number of interesting cocrystals; although if we compare the number of cocrystals to that known for more common coformer targets such as carboxylic acids and amides, it can be inferred that the area is still in its infancy (Table 4). For the purpose of this table, hydrates and solvates have been categorized separately from 'true cocrystals'.

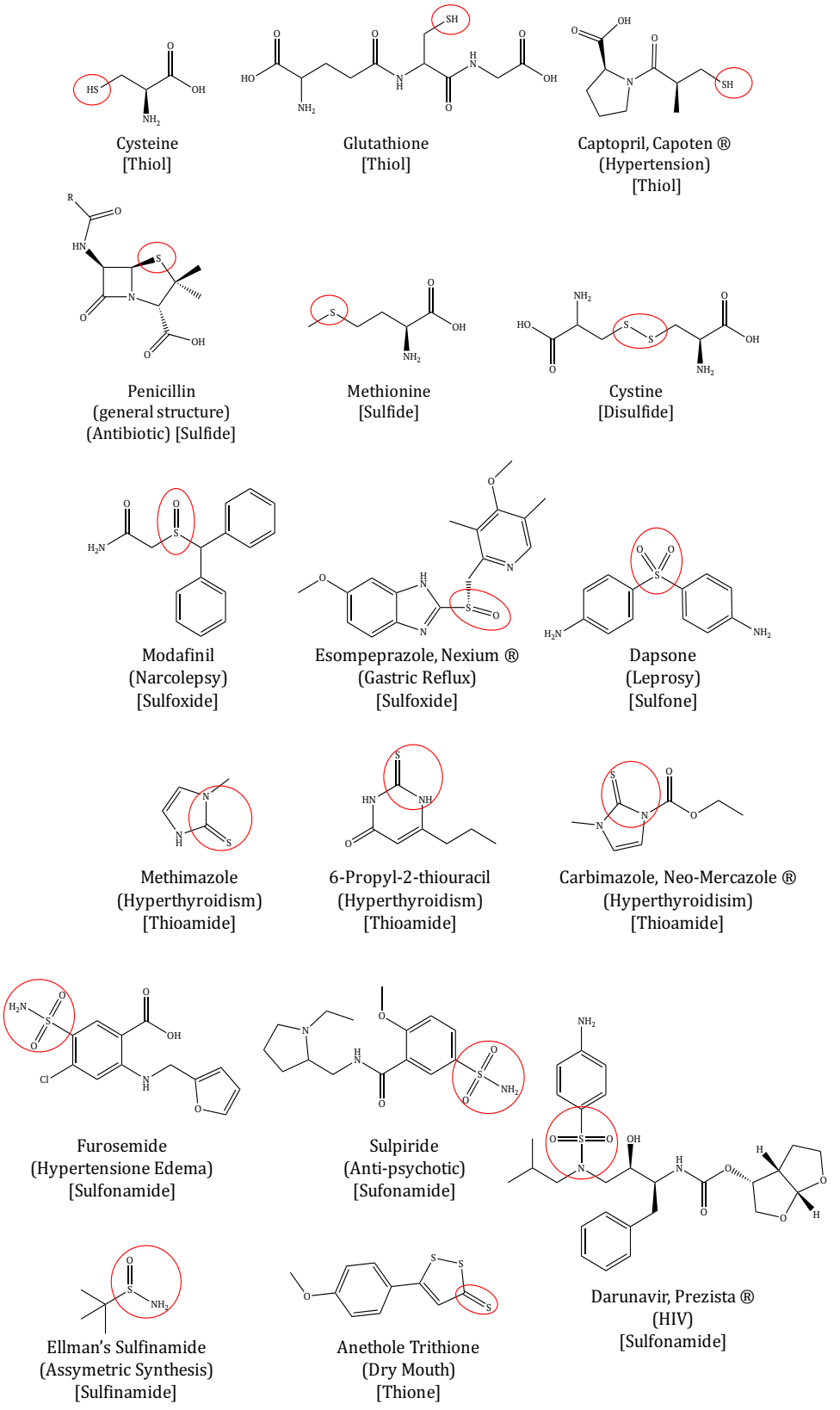


Figure 14: Examples of sulfur functional groups in pharmaceutical/biological materials (functional group circled in red).

Table 4: CSD[39] statistics for sulfur functional group cocrystals (Dec 2016).

Functional Group	Total Materials with Residues >1	True Cocrystals	Other (including solvates/hydrates)
Thiol	29	14	15
Sulfide	1057 (3650[a])	1057[a]	2593[a]
Disulfide	253	54	199
Sulfoxide	225 (1477[b])	75	150
Sulfone	411	78	333
Sulfinamide			
Primary	0	0	0
Secondary	42	1	41
Sulfonamide			
Primary	199	154	45
Secondary	481	135	346
Thioamide	153	72	81
Thione	21	3	18

a - Solvated and hydrated materials were excluded from second search due to high number of initial results which may have excluded some solvated cocrystals; b - 1477 materials including dimethylsulfoxide (DMSO) solvates, which were subsequently excluded from the search.

The CSD[39] was used to investigate the landscape of sulfur cocrystals already reported in the literature (Table 4). Motif analysis in Mercury[41] shows that the thiol, thione, sulfide and disulfide functional groups are not responsible for cocrystal formation, with the main hydrogen bonding interactions occurring from/to other functional groups in the materials.

A more diverse suite of intermolecular interactions can be observed from the oxidized sulfur functional groups, sulfoxide, sulfone, sulfinamide and sulfonamide. The sulfoxide moiety features heavily in the CSD,[39] particularly dimethylsulfoxide (DMSO) which features in 1252 multi-component crystals (which should be classified as solvates). Considerable research into the cocrystallization of sulfoxides has been performed within our research group, revealing both dibenzylsulfoxide (DBSO, **7**) and diphenylsulfoxide (DPSO, **8**) as potent hydrogen/halogen bond acceptors in cocrystallization. The sulfoxide (S=O) group is capable of hydrogen[80,81] and halogen[82] bonding with a range of donors, and is even capable of breaking the strong amide $R_2^2(8)$ dimer, a particularly persistent and dependable motif [Figure 15].

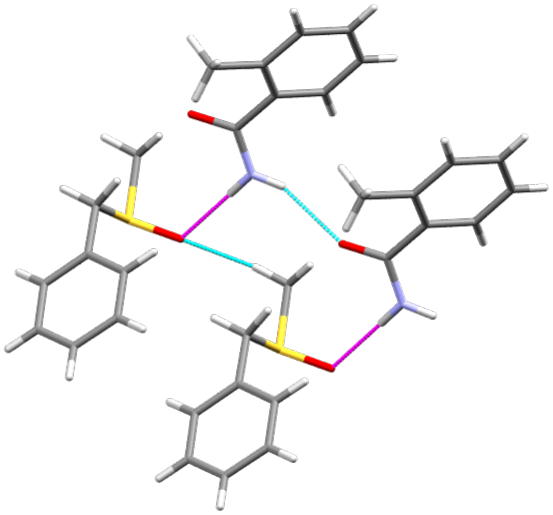


Figure 15: Cocrystallisation with DBSO [7] and 2-methylbenzamide [9] showing breakage of the amide dimer (one aromatic ring on DBSO removed for clarity, hydrogen bonds in magenta and cyan) [DALPOX].[80]

In contrast, in cocrystallization of diphenyl sulfoxide with 4-aminobenzoic acid [**2**] (ENAGIK)[80] the acid-acid $R_2^2(8)$ heterodimer is retained, with S=O···H-N hydrogen bonds occurring from the sulfoxide to the *para*-amino group instead, forming an $R_4^2(8)$ tetramer [Figure 16].

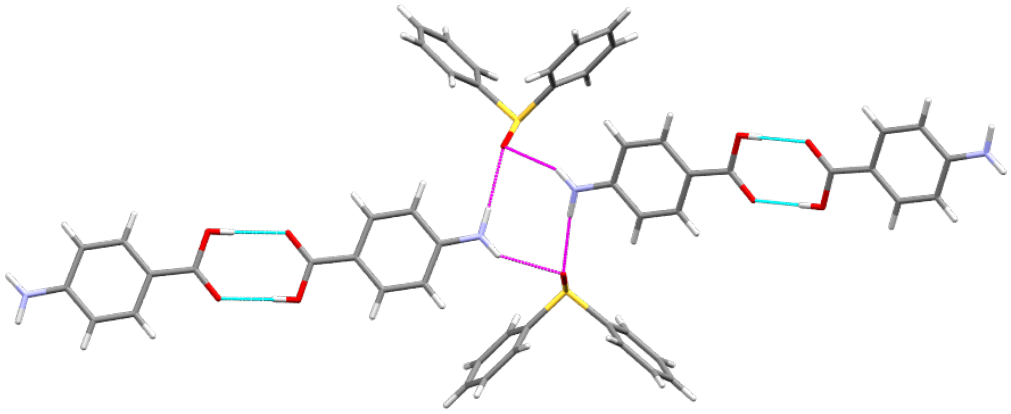


Figure 16: Retention of carboxylic acid dimer in cocrystallisation with diphenylsulfoxide [hydrogen bonds in magenta and cyan] (ENAGIK).[80]

The literature shows that the sulfoxide S=O can interact with the carboxylic acid, and amide functional groups in the assembly of multi-component crystals, accepting a hydrogen bond from the carboxylic acid O-H (73% frequency), or the amide/amine N-H (92% frequency) (Figures 17-19). Generally, Etter's rules are retained, with the best acceptor (sulfoxide) accepting the hydrogen bond from the best donor (acid/amide), and in the case of 3,5-dinitrosalicylic acid [**10**] (Figure 18), the S(6) ring is retained from the phenol to the carboxylic acid C=O.

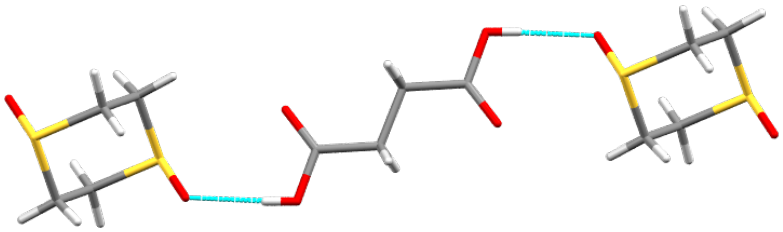


Figure 17: O-H···O=S interaction observed in 1:1 1,4-dithiane-1,4-dioxide - succinic acid cocrystal [11.12] (OGAHAF).[83]

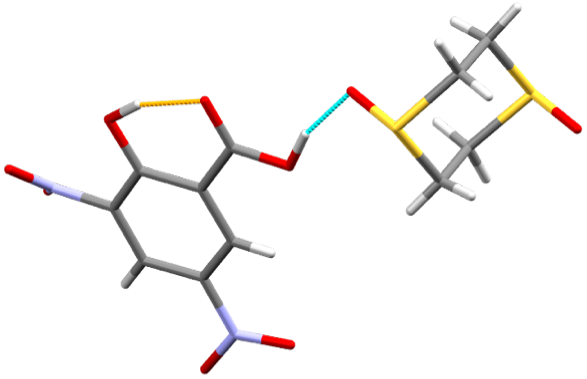


Figure 18: O-H···O=S interaction observed in sulfoxide cocrystal (OGAHEJ) [retention of the S(6) intramolecular ring in orange].[83]

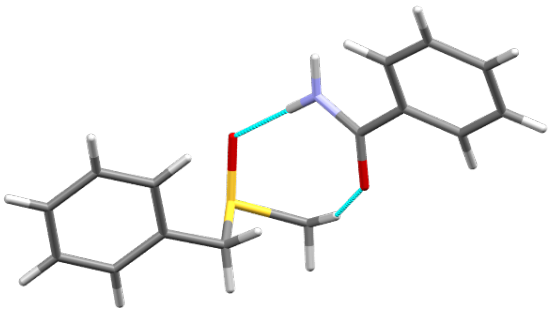


Figure 19: Sulfoxide-amide interactions observed in 1:1 cocrystal of benzamide - DBSO [13.8] (ENAHAD).[80]

The sulfone moiety is present in 78 cocrystals, of which 24 contain a hydrogen bond from the amine N-H to the sulfone S=O (N-H···O=S) forming a C(4) chain with a percentage frequency of 64%. Many of these observed interactions occur with aminated sulfone residues such as dapsone (4,4'-sulfonyldianiline, **14**) [Figure 20 and Figure 21], with the N-H···O=S interaction occurring between the molecules of **14**. Interestingly, the sulfone moiety does not participate in heteromeric 'cocrystal forming' interactions. All of the N-H···O=S observed are between the sulfone molecules themselves, with cocrystallization mediated by alternative hydrogen bonding interactions with the other component.

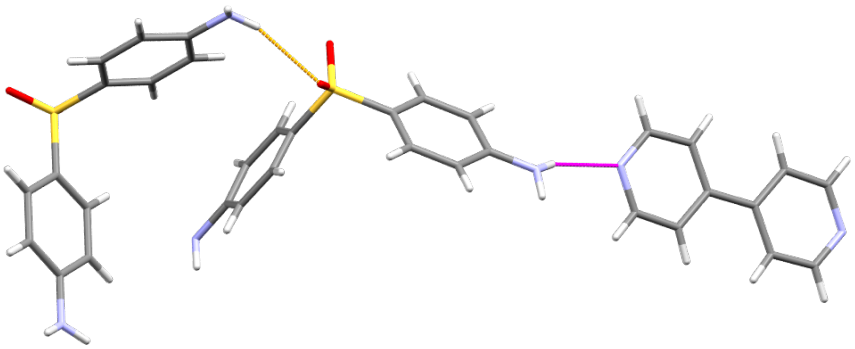


Figure 20: N-H···O=S interaction (orange) observed in 2:1 cocrystal of dapsone - 4,4'-bipyridyl [14.15] (KIGNEV).[84]

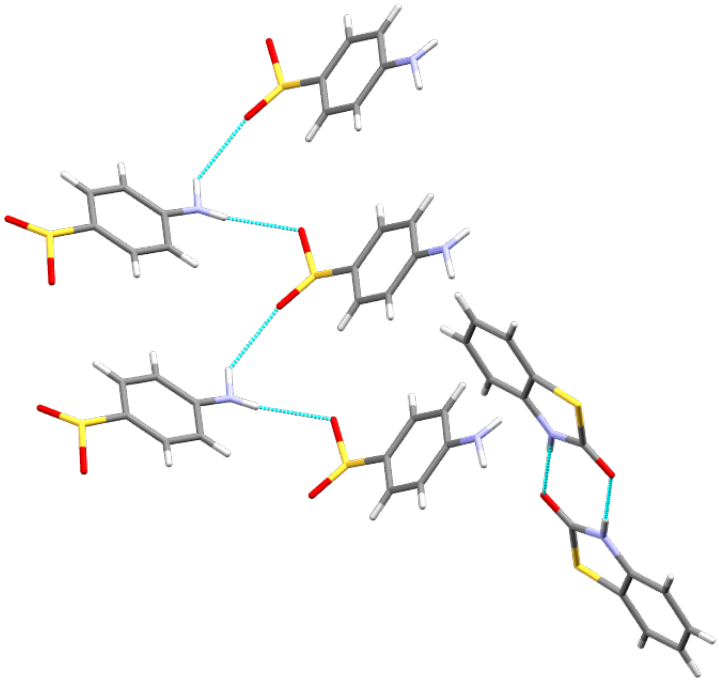


Figure 21: Dapsone C(4) chain observed in 1:1 drug-drug cocrystal of dapsone - 1,3-benzothiazole-2,3(H)-one [14.16] (VOHKUA).[85]

Other than N-H···O=S hydrogen bonds, the sulfone group can accept hydrogen bonds from less potent hydrogen bond donors such as aliphatic/aromatic C-H [Figure 22].

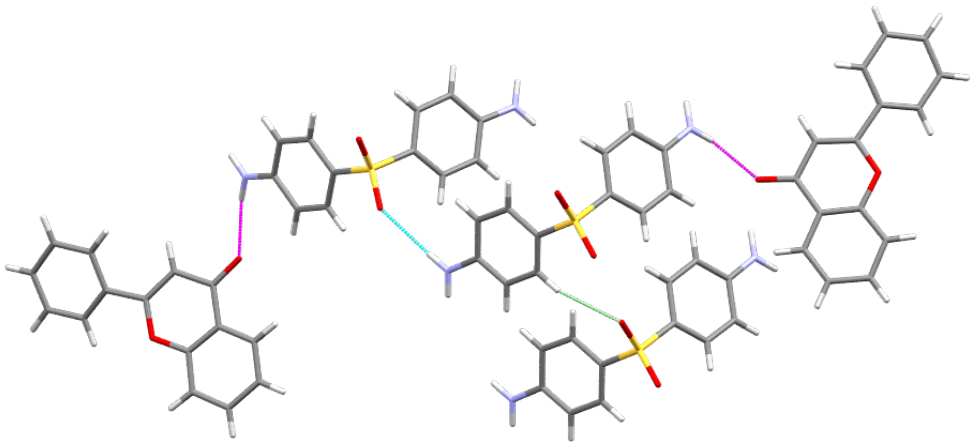


Figure 22: C-H···O=S [green] and N-H···O=C [magenta] interactions, in addition to the previously described chain (cyan) in 1:1 drug-drug cocrystal of dapsone - 2-phenyl-4H-chromen-4-one [14.17] (VOKHEK).[85]

The sulfone moiety accepts hydrogen bonds less readily than its sulfoxide counterpart, making it a less attractive functional group for cocrystallization in the absence of additional good hydrogen bond donors/acceptors also present on the molecule. The sulfone moiety was investigated previously within our research group as part of our work on sulfoxides, with no successful outcomes.[80–82] It was noted that the increased polarity of the sulfoxide group by comparison to the sulfone is a contributor to the promotion of intermolecular interactions and thus, cocrystallization.[82] This is in agreement with Hunter, the sulfone group has an assigned value of 6.3 in the table,[51] with the sulfoxide given 8.9 (9.0 being the highest in the table).

There is just a single example of a cocrystal with a sulfinamide-type group in the CSD,[39] a 1:1 cocrystal of 1,2-benzothiazol-3(2H)-one [**18**] with 1,2-benzothiazol-3(2H)-one-1-oxide [**19**], although one could debate as to whether this material should best be described as an amide or a sulfinamide. The primary hydrogen bonding motifs in this cocrystal are driven by the amide moiety, with the amide $R_2^2(8)$ dimer motif at the centre of the structure, capped by interactions to 'sulfinamide' N-H. The 'sulfinamide' S=O accepts a moderate hydrogen bond from a neighbouring C-H group at a distance of 2.5 Å [Figure 23]. This is an interesting example of the slight flexibility of Hunter's table and Etter's rules. In this case, preference is given to the formation of symmetrical bonding motifs (amide dimer) over the donation of the hydrogen bond from the strongest donor

(N-H) to the strongest acceptor (S=O). The S=O lies out of the plane of the ring, and so cannot allow for the formation of this symmetrical relationship.

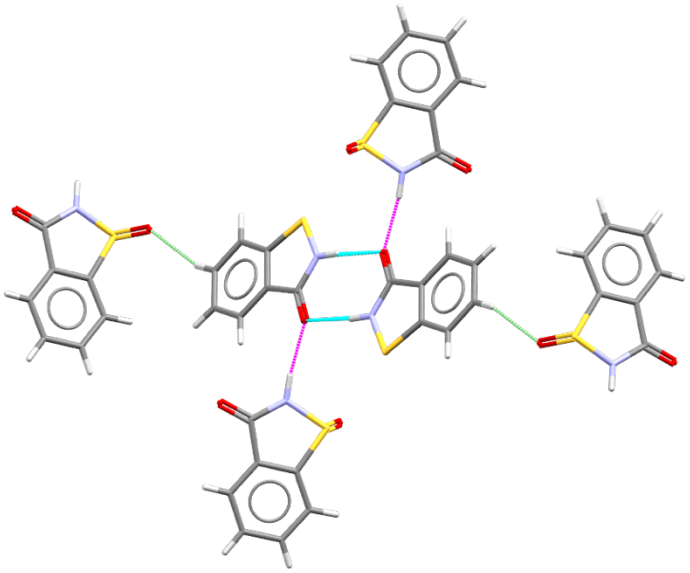


Figure 23: Hydrogen bonding in 1:1 cocrystal of 1,2-benzothiazol-3(2H)-one with 1,2-benzothiazol-3(2H)-one-1-oxide [18.19] (MUSPAS).[86]

The sulfonamide group has been the subject of many cocrystallization studies due to its presence in a broad range of active pharmaceutical materials.[87–90] The sulfonamide moiety itself is an attractive target for cocrystallization *via* hydrogen bonding, containing both a good acceptor (S=O), and a strong donor (N-H). Sulfonamides participate in a range of heteromeric synthons with various functional groups including (but not limited to) amides, carboxylic acids, amines, pyridyl compounds, *N*-oxides, alcohols, phosphine oxides, and polyethers.

Amongst this broad and varied set of cocrystals, some interesting motifs can be identified, firstly concerning the participation of the sulfonamide functionality in hydrogen bonded motifs with itself within the cocrystal structures. Two main homomeric motifs were identified in the sulfonamide cocrystal structures, the C(4) chain and the $R_2^2(8)$ dimer, with cocrystallization occurring as result of additional heteromeric motifs. The structure of celecoxib - valerolactam [**20.21**] (LIQQIM)[87] was published in 2014 by Nangia and coworkers as part of a series of celecoxib [**20**] cocrystals.[87] The structure contains a symmetrical sulfonamide $R_2^2(8)$ dimer at the centre of the structure, which is capped by the molecules of

valerolactam [**21**] (Figure 24). This homomeric dimer can be observed in other sulfonamide cocrystals also (19 examples in CSD).[39]

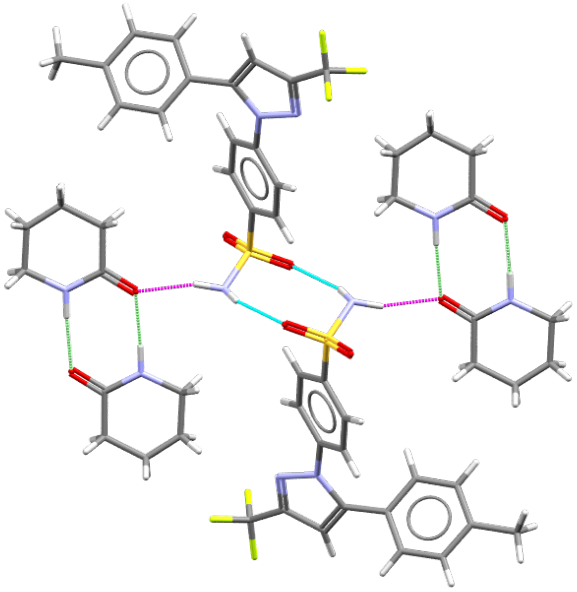


Figure 24: $R_2^2(8)$ dimer dimer motif observed in celecoxib - valerolactam cocrystal [20.21] [hydrogen bonds in magenta, green and cyan].[87]

The sulfonamide C(4) chain motif is present in 11 structures, including the 1:1 cocrystal of furosemide-nicotinamide [**22.1**], published by Ueto *et al*. in 2012.[91] In this case, the chain occurs between the sulfonamide groups, utilizing one of the two N-H atoms, while the other N-H participates in a heteromeric sulfonamide-amide interaction [Figure 25]. Similarly, in the case of the acetazolamide-2-pyridone cocrystal [**23.24**] published by Bolla and Nangia in 2016[92] the C(4) chain persists between the sulfonamide groups, with heteromeric interactions occurring at other parts of the molecules of **23** [Figure 26].

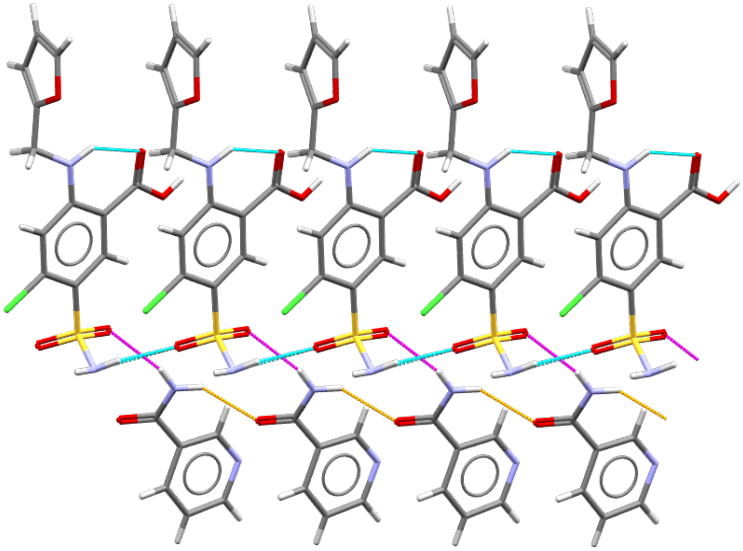


Figure 25: Homo and heteromeric sulfonamide interactions observed in 1:1 furosemide - nicotinamide cocrystal [22.1] (YASGOQ) [hydrogen bonds in magenta, orange and cyan].[91]

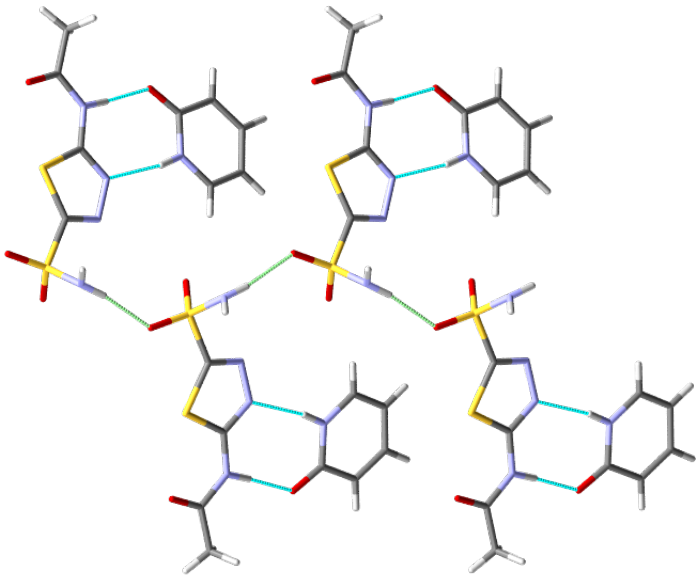


Figure 26: C(4) chain [green] observed in 1:1 cocrystal of acetazolamide - 2-pyridone [23.24] (MADSAO).[92]

There are several examples of heteromeric hydrogen bond donation from sulfonamide N-H to other functional groups [Table 5], including the pyridyl (Figure 27),[93–96] carbonyl (Figure 28),[97–100] ether (Figure 29),[98,101–104] *N*-oxide (Figure 30),[105–110] and to a lesser extent phosphine oxide functional groups (Figure 31).[111,112] The sulfonamide S=O participates in heteromeric hydrogen bond acceptor interactions to hydroxyls (Figure 32)[88,113,114] and amino groups (Figure 33).[87,89,90,114]

Table 5: Overview of heteromeric interactions observed in sulfonamide cocrystals.

Sulfonamide N-H Donor Interactions		
	No. of Cocrystal Examples	**% Frequency**
Pyridyl	41	45%
Carbonyl (any)	117	65%
Ethers (including clathrates)	23	26%
N-oxides	11	N/A[a]
Phosphine Oxides	2	N/A[a]
Sulfonamide S=O Acceptor Interaction		
	No. of Cocrystal Examples	**% Frequency**
Amino (N-H)	122	56%
Hydroxyl (O-H)	13	N/A[a]

a – Too few examples for % frequency to be statistically meaningful.

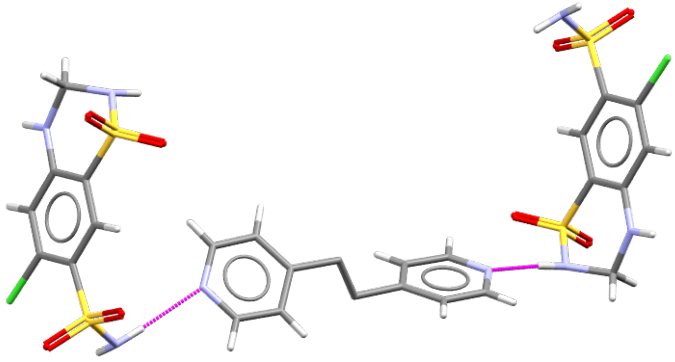


Figure 27: Sulfonamide N-H to pyridyl interactions in EGENUB.[115]

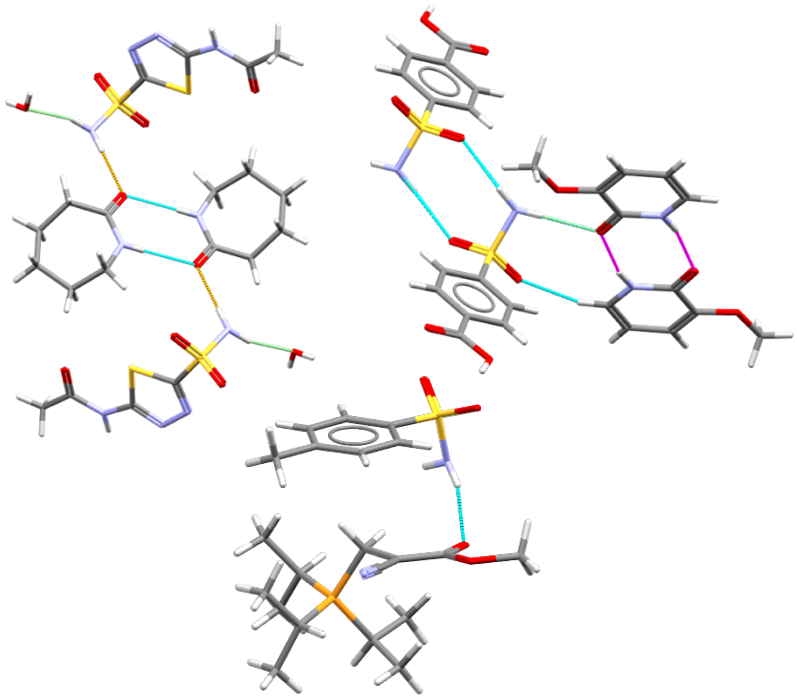


Figure 28: Sulfonamide N-H to carbonyl C=O interactions in MADSIW [top left] and FURYAU [top right, green] and NETDIZ [bottom].[89,92,116]

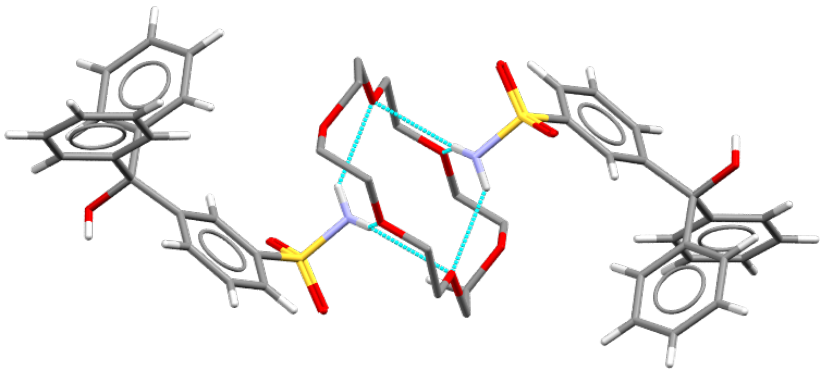


Figure 29: Sulfonamide N-H to ether clathrate in WOPCIO.[113]

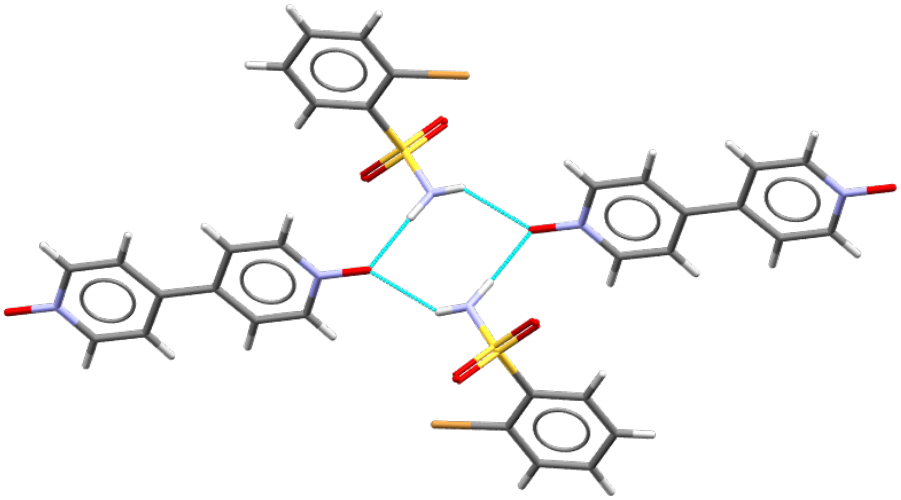


Figure 30: Sulfonamide N-H to *N*-oxide O-N hydrogen bonds in IWEQUX.[108]

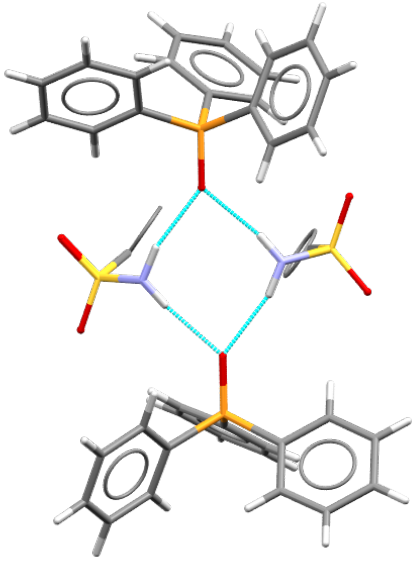


Figure 31: Sulfonamide N-H to phosphine oxide P=O hydrogen bonds in DATGIQ [aromatic rings removed for clarity].[112]

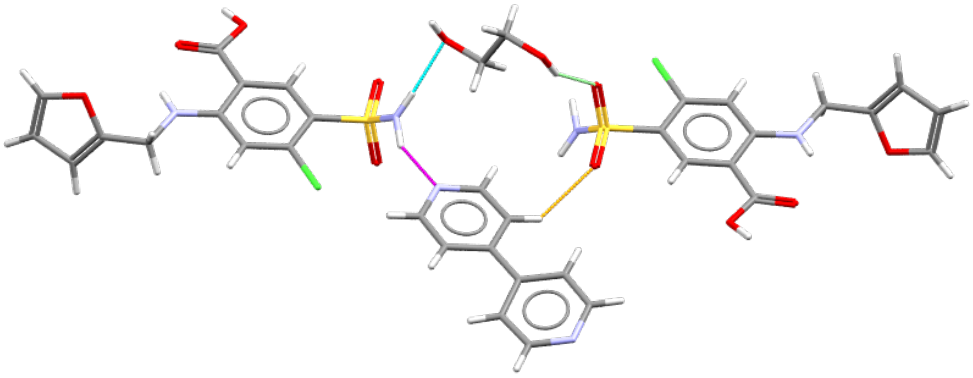


Figure 32: Hydroxyl O-H to sulfonamide S=O hydrogen bonds [green] in BOKHOA.[88]

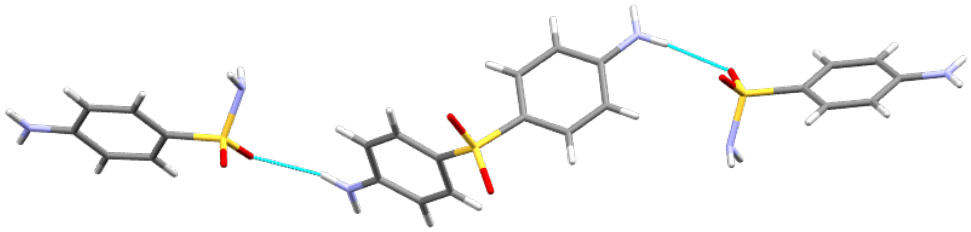


Figure 33: Amino N-H to sulfonamide S=O hydrogen bonds in VOHKAG, 4-aminobenzene sulfonamide - dapsone [25.14] cocrystal.[85]

The large variation in observed motifs for sulfonamide cocrystals provides an interesting basis for the assembly of multi-component materials, because the number of potential coformers is large. However, the small sample set available at this time would make it difficult to identify the robust synthons required for development of a library of suitable coformers for cocrystallization.

The thioamide functional group presents significantly less variation in the observed motifs. The primary hydrogen bonding interactions observed here are (a) clathrate structures incorporating donation of the thioamide N-H to an ether oxygen atom (24 examples, Figure 34),[117–119] (b) donation of hydrogen bonds from the N-H to an aromatic/cyclic nitrogen (12 examples, Figure 35)[120–123] and (c) halogen bonding interactions to the thioamide group (13 examples).[124] There are also 8 examples of interactions from the thioamide to a carbonyl C=O.[125–129] Interactions of the thioamide with dicarboxylic acids were recently published, showing significantly improved solubility profiles for the water insoluble drug ethionamide [**26**] (Figure 36).[129]

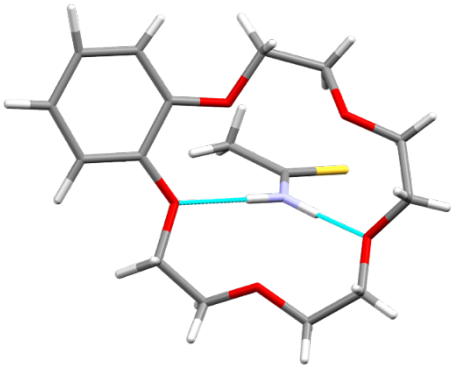


Figure 34: Thioamide N-H to ether clathrate stucture, FIXYEQ.[130]

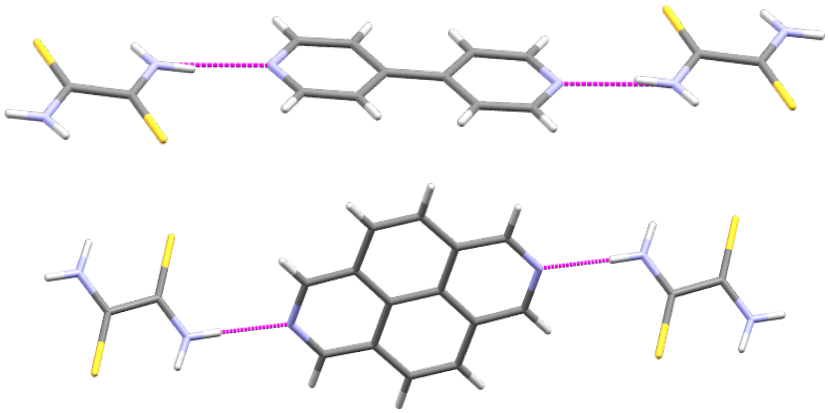


Figure 35: Thioamide N-H to pyridyl nitrogen interactions (magenta) in DOKGEQ (top), and DOKGUA (bottom).[123]

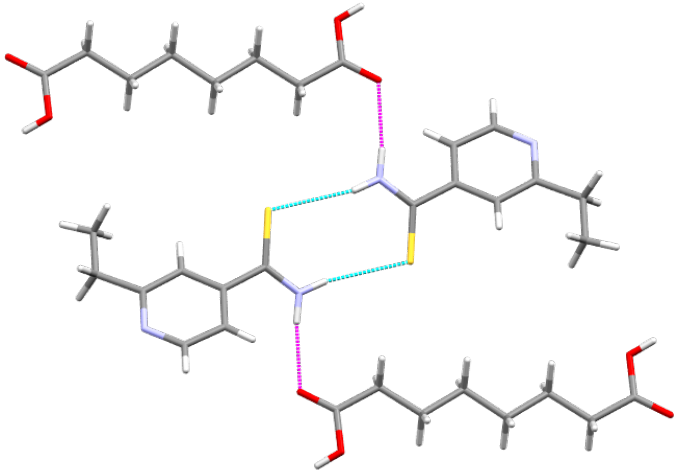


Figure 36: Thioamide N-H to carboxylic acid C=O (magenta) in 1:1 cocrystal of ethionamide - suberic acid [26.27] (NAKBIN).[129]

1.3 Objectives

Specific objectives will be outlined in each chapter going forward. The general focus of this work is in the utilisation of our current understanding of intermolecular interactions, particularly hydrogen bonding, incorporating Etter's rules and Hunter's table to design and investigate homomeric and heteromeric hydrogen bonding interactions in the solid state, with particular focus on the investigation of understudied functional groups, namely the primary aryl sulfinamide group and secondly, the use of this knowledge in the design and synthesis of cocrystals.

Chapter 1

References

1.4 References

(1) Desiraju, G. R. *J. Am. Chem. Soc.* **2013**, *135*, 9952–9967.

(2) Desiraju, G. R. *Agnew. Chem. Int. Ed.* **2007**, *46*, 8342–8356.

(3) Zukerman-Schpector, J.; Tiekink, E. R. T. *Z. fur Krist.* **2008**, *223*, 233–234.

(4) Banerjee, R.; Day, G. M.; Friščić, T.; Zhang, H. *CrystEngComm* **2016**, *18*, 3963–3967.

(5) Schmidt, G. M. J. *Pure Appl. Chem.* **1971**, *27*, 647–678.

(6) Bragg, W. H. *Proc. Phys. Soc. London* **1921**, *34*, 33–50.

(7) Desiraju, G. R. *Crystal Engineering. The Design of Organic Solids*; Elsevier: Amsterdam, 1989.

(8) Steed, J. W.; Atwood, J. L. *Supramolecular Chemistry, 2nd Ed.*, 2nd ed.; Wiley: UK, 2009.

(9) Lehn, J. M. *Supramolecular Chemistry: Concepts and Perspectives*; VCH, Weinheim, 1995.

(10) Desiraju, G. R. *Agnew. Chem. Int. Ed. Engl.* **1995**, *34*, 2311–2327.

(11) Beatty, A. M.; Helfrich, B. A.; Nieuwenhuyzen, M. *Cryst. Growth Des.* **2003**, *3*, 159–165.

(12) Aakeröy, C. B.; Salmon, D. J. *CrystEngComm* **2005**, *7*, 439–448.

(13) Childs, S. L.; Chyall, L. J.; Dunlap, J. T.; Smolenskaya, V. N.; Stahly, B. C.; Stahly, G. P. *J. Am. Chem. Soc.* **2004**, *126*, 13335–13342.

(14) Lipper, R. B. *Mod. Drug Discov.* **1999**, *2*, 55.

(15) Liu, R. *Water-Insoluble Drug Formulation*, 2nd ed.; CRC Press, Taylor and Francis Group, Boca Raton, FL, 2008.

(16) Schwiebert, K. E.; Chin, D. N.; MacDonald, J. C.; Whitesides, G. M. *J. Am. Chem. Soc.* **1996**, *118*, 4018–4029.

(17) Maddox, J. *Nature* **1988**, *335*, 201.

(18) Gavezzotti, A. *Acc. Chem. Res* **1994**, *27*, 309–314.

(19) Woodley, S. M.; Catlow, R. *Nat. Mater.* **2008**, *7*, 937–946.

(20) David, W. I. F.; Shankland, K.; Van De Streek, J.; Pidcock, E.; Motherwell, W. D. S.; Cole, J. C. *J. Appl. Crystallogr.* **2006**, *39*, 910–915.

(21) Oganov, A. R.; Glass, C. W. *J. Chem. Phys.* **2006**, *124*, 244704–244715.

(22) Wu, S. Q.; Ji, M.; Wang, C. Z.; Nguyen, M. C.; Zhao, X.; Umemoto, K.;

Wentzcovitch, R. M.; Ho, K. M. *J. Phys. Condens. Matter* **2014**, *26*, 35402–35407.

(23) Day, G. M. *Crystallogr. Rev.* **2011**, *17*, 3–52.

(24) Thompson, H. P. G.; Day, G. M. *Chem. Sci.* **2014**, 1–11.

(25) Wicker, J. G. P.; Cooper, R. I. *CrystEngComm* **2015**, *17*, 1927–1934.

(26) Lommerse, J. P. M.; Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Mooij, W. T. M.; Price, S. L.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. *Acta Crystallogr. Sect. B Struct. Sci.* **2000**, *56*, 697–714.

(27) Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Dzyabchenko, A.; Erk, P.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Lommerse, J. P. M.; Mooij, W. T. M.; Price, S. L.; Scheraga, H.; Schweizer, B.; Schmidt, M. U.; Van Eijck, B. P.; Verwer, P.; Williams, D. E. *Acta Crystallogr. Sect. B* **2002**, *58* (4), 647–661.

(28) Day, G. M.; Motherwell, W. D. S.; Ammon, H. L.; Boerrigter, S. X. M.; Della Valle, R. G.; Venuti, E.; Dzyabchenko, A.; Dunitz, J. D.; Schweizer, B.; Van Eijck, B. P.; Erk, P.; Facelli, J. C.; Bazterra, V. E.; Ferraro, M. B.; Hofmann, D. W. M.; Leusen, F. J. J.; Liang, C.; Pantelides, C. C.; Karamertzanis, P. G.; Price, S. L.; Lewis, T. C.; Nowell, H.; Torrisi, A.; Scheraga, H. A.; Arnautova, Y. A.; Schmidt, M. U.; Verwer, P. *Acta Crystallogr. Sect. B* **2005**, *61*, 511–527.

(29) Day, G. M.; Cooper, T. G.; Cruz-Cabeza, A. J.; Hejczyk, K. E.; Ammon, H. L.; Boerrigter, S. X. M.; Tan, J. S.; Della Valle, R. G.; Venuti, E.; Jose, J.; Gadre, S. R.; Desiraju, G. R.; Thakur, T. S.; Van Eijck, B. P.; Facelli, J. C.; Bazterra, V. E.; Ferraro, M. B.; Hofmann, D. W. M.; Neumann, M. A.; Leusen, F. J. J.; Kendrick, J.; Price, S. L.; Misquitta, A. J.; Karamertzanis, P. G.; Welch, G. W. A.; Scheraga, H. A.; Arnautova, Y. A.; Schmidt, M. U.; Van De Streek, J.; Wolf, A. K.; Schweizer, B. *Acta Crystallogr. Sect. B* **2009**, *65*, 107–125.

(30) Bardwell, D. A.; Adjiman, C. S.; Arnautova, Y. A.; Bartashevich, E.; Boerrigter, S. X. M.; Braun, D. E.; Cruz-Cabeza, A. J.; Day, G. M.; Della Valle, R. G.; Desiraju, G. R.; Van Eijck, B. P.; Facelli, J. C.; Ferraro, M. B.; Grillo, D.; Habgood, M.; Hofmann, D. W. M.; Hofmann, F.; Jose, K. V. J.; Karamertzanis, P. G.; Kazantsev, A. V.; Kendrick, J.; Kuleshova, L. N.; Leusen, F. J. J.; Maleev,

A. V.; Misquitta, A. J.; Mohamed, S.; Needs, R. J.; Neumann, M. A.; Nikylov, D.; Orendt, A. M.; Pal, R.; Pantelides, C. C.; Pickard, C. J.; Price, L. S.; Price, S. L.; Scheraga, H. A.; Van De Streek, J.; Thakur, T. S.; Tiwari, S.; Venuti, E.; Zhitkov, I. K. *Acta Crystallogr. Sect. B* **2011**, *67*, 535–551.

(31) Reilly, A. M.; Cooper, R. I.; Adjiman, C. S.; Bhattacharya, S.; Boese, A. D.; Brandenburg, J. G.; Bygrave, P. J.; Bylsma, R.; Campbell, J. E.; Car, R.; Case, D. H.; Chadha, R.; Cole, J. C.; Cosburn, K.; Cuppen, H. M.; Curtis, F.; Day, G. M.; DiStasio, R. A.; Dzyabchenko, A.; Van Eijck, B. P.; Elking, D. M.; Van Den Ende, J. A.; Facelli, J. C.; Ferraro, M. B.; Fusti-Molnar, L.; Gatsiou, C. A.; Gee, T. S.; De Gelder, R.; Ghiringhelli, L. M.; Goto, H.; Grimme, S.; Guo, R.; Hofmann, D. W. M.; Hoja, J.; Hylton, R. K.; Iuzzolino, L.; Jankiewicz, W.; De Jong, D. T.; Kendrick, J.; De Klerk, N. J. J.; Ko, H. Y.; Kuleshova, L. N.; Li, X.; Lohani, S.; Leusen, F. J. J.; Lund, A. M.; Lv, J.; Ma, Y.; Marom, N.; Masunov, A. E.; McCabe, P.; McMahon, D. P.; Meekes, H.; Metz, M. P.; Misquitta, A. J.; Mohamed, S.; Monserrat, B.; Needs, R. J.; Neumann, M. A.; Nyman, J.; Obata, S.; Oberhofer, H.; Oganov, A. R.; Orendt, A. M.; Pagola, G. I.; Pantelides, C. C.; Pickard, C. J.; Podeszwa, R.; Price, L. S.; Price, S. L.; Pulido, A.; Read, M. G.; Reuter, K.; Schneider, E.; Schober, C.; Shields, G. P.; Singh, P.; Sugden, I. J.; Szalewicz, K.; Taylor, C. R.; Tkatchenko, A.; Tuckerman, M. E.; Vacarro, F.; Vasileiadis, M.; Vazquez-Mayagoitia, A.; Vogt, L.; Wang, Y.; Watson, R. E.; De Wijs, G. A.; Yang, J.; Zhu, Q.; Groom, C. R. *Acta Crystallogr. Sect. B* **2016**, *72*, 439–459.

(32) Desiraju, G. R. *Nature* **2001**, *412*, 397–400.

(33) Desiraju, G. R. *Acc. Chem. Res.* **2002**, *35*, 565–573.

(34) Arunan, E.; Desiraju, G. R.; Klein, R. A.; Sadlej, J.; Scheiner, S.; Alkorta, I.; Clary, D. C.; Crabtree, R. H.; Dannenberg, J. J.; Hobza, P.; Kjaergaard, H. G.; Legon, A. C.; Mennucci, B.; Nesbitt, D. J. *Pure Appl. Chem.* **2011**, *83*, 1637–1641.

(35) Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press: Oxford, 1997.

(36) Steiner, T. *Angew. Chem. Int. Ed.* **2002**, *41*, 48–76.

(37) Shahi, A.; Arunan, E. *J. Chem. Sci.* **2016**, *128*, 1571–1577.

(38) Wood, P. A.; Allen, F. H.; Pidcock, E. *CrystEngComm* **2009**, *11*, 1563–1571.

(39) R., G. C.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. *Acta Crystallogr. Sect. B* **2016**, *72*, 171–179.

(40) *Diam. version 3.1, Cryst. Impact,* **2006**, 1251.

(41) Macrae, C. F.; Bruno, I. J.; Chisholm, J. A.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; van de Streek, J.; Wood, P. A. J. *J. Appl. Crystallogr.* **2008**, *41*, 466–470.

(42) Spek, A. L. *J. Appl. Crystallogr.* **2003**, *36*, 7–13.

(43) Sheldrick, G. M. *Acta Crystallogr. Sect. A* **2008**, *64*, 112–122.

(44) Etter, M. C. *J. Am. Chem. Soc.* **1982**, *502*, 1095–1096.

(45) Etter, M. C. *J. Phys. Chem.* **1991**, *95*, 4601–4610.

(46) Leiserowitz, L.; Nader, F. *Acta Crystallogr. Sect. B* **1977**, 2719–2733.

(47) Leiserowitz, L.; Schmidt, G. M. J. *J. Chem. Soc.* **1969**, *500*, 2372–2382.

(48) Donohoe, J. *J. Phys. Chem.* **1952**, *56*, 502–510.

(49) Lai, T. F.; Marsh, R. E. *Acta Crystallogr.* **1967**, *22*, 885–893.

(50) Lin, C.-T.; Siew, P.-Y.; Byrn, S. R. *J. Chem. Soc., Perkin Trans. 2* **1978**, *10*, 957–962.

(51) Hunter, C. A. *Angew. Chemie - Int. Ed.* **2004**, *43*, 5310–5324.

(52) Abraham, M. H.; Platts, J. A. *J. Org. Chem.* **2001**, *66*, 3484–3491.

(53) Wohler, F. *Ann. Chemie Pharm.* **1844**, *51*, 145–163.

(54) Sakurai, T. *Acta Crystallogr.* **1965**, *19*, 320–330.

(55) Desiraju, G. R. *CrystEngComm* **2003**, *5*, 466.

(56) Dunitz, J. D. *CrystEngComm* **2003**, *5*, 506.

(57) *Food Drug Adm.* **2011**, *December*, 76 FR 75551.

(58) Aitipamula, S.; Banerjee, R.; Bansal, A. K.; Biradha, K.; Cheney, M. L.; Choudhury, A. R.; Desiraju, G. R.; Dikundwar, A. G.; Dubey, R.; Duggirala, N.; Ghogale, P. P.; Ghosh, S.; Goswami, P. K.; Goud, N. R.; Jetti, R. R. K. R.; Karpinski, P.; Kaushik, P.; Kumar, D.; Kumar, V.; Moulton, B.; Mukherjee, A.; Mukherjee, G.; Myerson, A. S.; Puri, V.; Ramanan, A.; Rajamannar, T.; Reddy, C. M.; Rodriguez-Hornedo, N.; Rogers, R. D.; Row, T. N. G.; Sanphui, P.; Shan, N.; Shete, G.; Singh, A.; Sun, C. C.; Swift, J. A.; Thaimattam, R.; Thakur, T. S.; Kumar Thaper, R.; Thomas, S. P.; Tothadi, S.; Vangala, V. R.; Variankaval, N.; Vishweshwar, P.; Weyna, D. R.; Zaworotko, M. J. *Cryst. Growth Des.* **2012**, *12*, 2147–2152.

(59) *Food Drug Adm.* **2013**, *April*, 78 FR 24754.

(60) *Food Drug Adm.* **2016**, *August*, 81 FR 54808.

(61) Grothe, E.; Meekes, H.; Vlieg, E.; Ter Horst, J. H.; De Gelder, R. *Cryst. Growth Des.* **2016**, *16*, 3237–3243.

(62) Wood, P. A.; Feeder, N.; Furlow, M.; Galek, P. T. A.; Groom, C. R.; Pidcock, E. *CrystEngComm* **2014**, *16*, 5839–5848.

(63) Bien, J. T.; Deshpande, P. P.; Dimarco, J. D.; Gougoutas, J. Z.; Grosso, J. A.; Chiajen Lai, H.; Lobinger, A. A.; Nirschl, A. A.; Ramakrishnan, S.; Riebel, P.; Singh, J.; Wang, C. Co-crystal of Dapagliflozin EP 2069374 A1, 2009.

(64) Remenar, J. F.; Morissette, S. L.; Peterson, M. L.; Moulton, B.; MacPhee, J. M.; Guzmán, H. R.; Almarsson, Ö. *J. Am. Chem. Soc.* **2003**, *125*, 8456–8457.

(65) Wang, Z. Z.; Chen, J. M.; Lu, T. B. *Cryst. Growth Des.* **2012**, *12*, 4562–4566.

(66) Karki, S.; Friščić, T.; Fabián, L.; Laity, P. R.; Day, G. M.; Jones, W. *Adv. Mater.* **2009**, *21*, 3905–3909.

(67) Grifasi, F.; Chierotti, M. R.; Gaglioti, K.; Gobetto, R.; Maini, L.; Braga, D.; Dichiarante, E.; Curzi, M. *Cryst. Growth Des.* **2015**, *15*, 1939–1948.

(68) Perlovich, G. L. *CrystEngComm* **2015**, *17*, 7019–7028.

(69) Qiao, N.; Li, M.; Schlindwein, W.; Malek, N.; Davies, A.; Trappitt, G. *Int. J. Pharm.* **2011**, *419*, 1–11.

(70) Fukte, S. R.; Wagh, M. P.; Rawat, S. *Int. J. Pharm. Pharm. Sci.* **2014**, *6*, 9–14.

(71) Wells, A. F. *Structural Inorganic Chemistry*, 3rd ed.; Clarendon Press: Oxford, 1962.

(72) Hamilton, W. C.; Ibers, J. A. In *Hydrogen Bonding in Solids*; W. A. Benjamin, Inc.: New York, 1968.

(73) Kuleshova, L. N.; Zorky, P. M. *Acta Crystallogr. Sect. B* **1980**, *36*, 2113–2115.

(74) Etter, M. C. *Acc. Chem. Res.* **1990**, *23*, 120–126.

(75) Etter, M. C. *Isr. J. Chem.* **1985**, *25*, 312–319.

(76) Bernstein, J.; Etter, M. C.; Macdonald, J. C. *J. Chem. Soc., Perkin Trans. 2* **1990**, 695–698.

(77) Etter, M. C.; MacDonald, J. C.; Bernstein, J. *Acta Crystallogr. Sect. B* **1990**, *46*, 256–262.

(78) Bernstein, J.; Davis, R. E.; Shimoni, L.; Chang, N. L. *Angew. Chemie Int. Ed.*

1995, *34*, 1555–1573.

(79) Feng, M.; Tang, B.; Liang, S. H.; Jiang, X. *Curr. Top Med. Chem.* **2016**, *16*, 1200–1216.

(80) Eccles, K. S.; Elcoate, C. J.; Stokes, S. P.; Maguire, A. R.; Lawrence, S. E. *Cryst. Growth Des.* **2010**, *10*, 4243–4245.

(81) Eccles, K. S.; Elcoate, C. J.; Maguire, A. R.; Lawrence, S. E. *Cryst. Growth Des.* **2011**, *11*, 4433–4439.

(82) Eccles, K. S.; Morrison, R. E.; Stokes, S. P.; O'Mahony, G. E.; Hayes, J. A.; Kelly, D. M.; O'Boyle, N. M.; Fábián, L.; Moynihan, H. A.; Maguire, A. R.; Lawrence, S. E. *Cryst. Growth Des.* **2012**, *12*, 2969–2977.

(83) Kumar, V. S.; Nangia, A.; Katz, A. K.; Carrell, H. L. *Cryst. Growth Des.* **2002**, *2*, 313–318.

(84) Martins, I.; Martins, M.; Fernandes, A.; Andre, V.; Teresa Duarte, M. *CrystEngComm* **2013**, *15*, 8173–8179.

(85) Jiang, L.; Huang, Y.; Zhang, Q.; He, H.; Xu, Y.; Mei, X. *Cryst. Growth Des.* **2014**, *14*, 4562–4573.

(86) Griffith, D. M.; Haughey, A.; Chahal, S.; Müller-Bunz, H.; Marmion, C. J. *Inorganica Chim. Acta* **2010**, *363*, 2333–2337.

(87) Bolla, G.; Mittapalli, S.; Nangia, A. *CrystEngComm* **2014**, *16*, 24–27.

(88) Srirambhatla, V. K.; Kraft, A.; Watt, S.; Powell, A. V. *CrystEngComm* **2014**, *16*, 9979–9982.

(89) Bolla, G.; Nangia, A. *Chem. Commun.* **2015**, *51*, 15578–15581.

(90) Bolla, G.; Mittapalli, S.; Nangia, A. *IUCrJ* **2015**, *2*, 389–401.

(91) Ueto, T.; Takata, N.; Muroyama, N.; Nedu, A.; Sasaki, A.; Tanida, S.; Terada, K. *Cryst. Growth Des.* **2012**, *12*, 485–494.

(92) Bolla, G.; Nangia, A. *IUCrJ* **2016**, *3*, 152–160.

(93) Elacqua, E.; Bucar, D. K.; Henry, R. F.; Zhang, G. G. Z.; MacGillivray, L. R. *Cryst. Growth Des.* **2013**, *13*, 393–403.

(94) Goud, N. R.; Khan, R. A.; Nangia, A. *CrystEngComm* **2014**, *16*, 5859–5869.

(95) Arman, H. D.; Kaulgud, T.; Tiekink, E. R. T. *Z. fur Krist.* **2014**, *229*, 394–404.

(96) Domingos, S.; Fernandes, A.; Duarte, M. T.; Piedade, M. F. M. *Cryst. Growth Des.* **2016**, *16*, 1879–1892.

(97) Caira, M. R. *J. Crystallogr. Spectrosc. Res.* **1991**, *21*, 641–648.

(98) Bock, H.; Nagel, N.; Eller, P. *Z. Naturforsch., B* **1999**, *54*, 501.

(99) Ghosh, S.; Bag, P. P.; Reddy, C. M. *Cryst. Growth Des.* **2011**, *11*, 3489–3503.

(100) Smith, G.; Wermuth, U. D. *Acta Crystallogr. Sect. E Struct. Reports Online* **2012**, *68*, o1649–o1650.

(101) Simonov, Y. A.; Malinowskii, T. I.; Ganin, E. V; Kotlyar, S. A.; Bocelli, G.; Calestani, G.; Rizzoli, C. *J. Incl. Phenom. Mol. Recognit. Chem.* **1990**, *8*, 349–361.

(102) Caira, M. R.; Mohamed, R. *Acta Crystallogr. Sect. B* **1993**, *49*, 760–768.

(103) Simonov, Y. A.; Fonari, M. S.; Duca, G. G.; Gonta, M. V.; Ganin, E. V.; Yavolovskii, A. A.; Gdaniec, M.; Lipkowski, J. *Tetrahedron* **2005**, *61*, 6596–6601.

(104) Paulidou, A.; Maffeo, D.; Yannakopoulou, K.; Mavridis, I. M. *CrystEngComm* **2010**, *12*, 517–525.

(105) Hamann, T.; Henschel, D.; Lange, I.; Moers, O.; Blaschette, A.; Jones, P. G. *Z. Naturforsch., B* **2002**, *57*, 1051.

(106) Saha, B. K.; Banerjee, R.; Nangia, A.; Desiraju, G. R. *Acta Crystallogr. Sect. E Struct. Reports Online* **2006**, *62*, o2283–o2284.

(107) Babu, N. J.; Reddy, L. S.; Nangia, A. *Mol. Pharm.* **2007**, *4*, 417–434.

(108) Rajesh Goud, N.; Jagadeesh Babu, N.; Nangia, A. *Cryst. Growth Des.* **2011**, *11*, 1930–1939.

(109) Rybarczyk-Pirek, A. J.; Zukomska-Rogala, M.; Wojtulewski, S.; Palusiak, M. *Cryst. Growth Des.* **2015**, *15*, 5802–5815.

(110) Puttreddy, R.; Jurček, O.; Bhowmik, S.; Mäkelä, T.; Rissanen, K. *Chem. Commun.* **2016**, *52*, 2338–2341.

(111) Ferguson, G.; Glidewell, C. *J. Chem. Soc. Perkin Trans. 2* **1988**, *12*, 2129–2132.

(112) Croker, D. M.; Foreman, M. E.; Hogan, B. N.; Maguire, N. M.; Elcoate, C. J.; Hodnett, B. K.; Maguire, A. R.; Rasmuson, Å. C.; Lawrence, S. E. *Cryst. Growth Des.* **2012**, *12*, 869–875.

(113) Fonari, M. S.; Simonov, Y. A.; Wang, W.-J.; Tang, S.-W.; Ganin, E. V. *CrystEngComm* **2009**, *11*, 94–101.

(114) Arenas-García, J. I.; Herrera-Ruiz, D.; Mondragón-Vásquez, K.; Morales-Rojas, H.; Höpfl, H. *Cryst. Growth Des.* **2012**, *12*, 811–824.

(115) Wang, J. R.; Ye, C.; Mei, X. *CrystEngComm* **2014**, *16*, 6996–7003.

(116) Golobov, Y. G.; Pinchuk, V. A.; Thonnessen, H.; Jones, P. G.; Schmutzler, R. *Phosphorous, Sulfur Silicon Relat. Elem.* **1996**, *115*, 19.

(117) Bocelli, G. *Org. Biomol. Chem.* **2005**, *3*, 3054–3058.

(118) Fonari, M. S.; Ganin, E. V.; Tang, S. W.; Wang, W. J.; Simonov, Y. A. *J. Mol. Struct.* **2007**, *826*, 89–95.

(119) Fonari, M. S.; Simonov, Y. A.; Bocelli, G.; Ganin, E. V.; Wang, W.-J.; Pluzhnik-Gladyr, S. M.; Tkachyuk, V. V.; Kotlyar, S. A.; Kamalov, G. L. *J. Incl. Phenom. Macrocycl. Chem.* **2008**, *61*, 367–375.

(120) Saenger, W.; Suck, D. *J. Mol. Biol;* **1971**, *60*, 87–99.

(121) Widlicka, D. W.; Wong, E. H.; Weisman, G. R.; Sommer, R. D.; Incarvito, C. D.; Rheingold, A. L. *Inorganica Chim. Acta* **2002**, *341*, 45–53.

(122) Olszewska, T.; Gdaniec, M.; Połoński, T. *J. Org. Chem.* **2004**, *69*, 1248–1255.

(123) Piotrkowska, B.; Wasilewska, A.; Gdaniec, M.; Połoński, T. *CrystEngComm* **2008**, *10*, 1421–1428.

(124) Eccles, K. S.; Morrison, R. E.; Sinha, A. S.; Maguire, A. R.; Lawrence, S. E. *Cryst. Growth Des.* **2015**, *15*, 3442–3451.

(125) Sharanin, Y. A.; Shestaopalov, A. M.; Nesterov, V. N.; Litvinov, V. P.; Mortikov, V. Y.; Promonenkov, V. K.; Shklover, V. E.; Struchkov, Y. T. *Chem. Heterocycl. Compd. Russ.* **1987**, 1377.

(126) Mizuguchi, J.; Rihs, G. *Acta Crystallogr., Sect. C Cryst. Struct. Commun.* **1992**, *48*, 1279–1283.

(127) Inoue, Y.; Kanbara, T.; Yamamoto, T. *Tetrahedron Lett.* **2004**, *45*, 4603–4606.

(128) Xu, L.-L.; Chen, J.-M.; Yan, Y.; Lu, T.-B. *Cryst. Growth Des.* **2012**, *12*, 6004–6011.

(129) Mannava, M. K. C.; Suresh, K.; Nangia, A. *Cryst. Growth Des.* **2016**, *16*, 1591–1598.

(130) Watson, W. H.; Jain, P. C. *J. Incl. Phenom. Mol. Recog. Chem* **1986**, *4*, 397.

Chapter 2

Synthesis and Solid State Characterisation of Primary Aryl Sulfinamides

2.1 Introduction to Primary Sulfinamides

The sulfinamide moiety can be considered as an interesting chiral analogue of the traditional, carbon-based amide. Where amides find extensive use in solid state chemistry, primary sulfinamides have captured little focus from this perspective, partly due to the far more challenging synthesis of the primary sulfinamide. Apart from one small cocrystallization study carried out in our lab in 2013[1] that utilised a commercially available, enantiopure-primary sulfinamide, there has been no research conducted solely to elucidate the crystal landscape of these materials. Structurally-related sulfonamides and primary amides have 486 and 2546 crystal structures reported in the CSD[2] respectively, and secondary sulfinamides account for 378 reported structures. Sulfinamide chirality arises from the four substituents on the sulfur atom, including one lone pair, accompanied by a trigonal orientation at the nitrogen, whereas the amide functional group is completely planar due to sp^2 hybridisation (Figure 1).

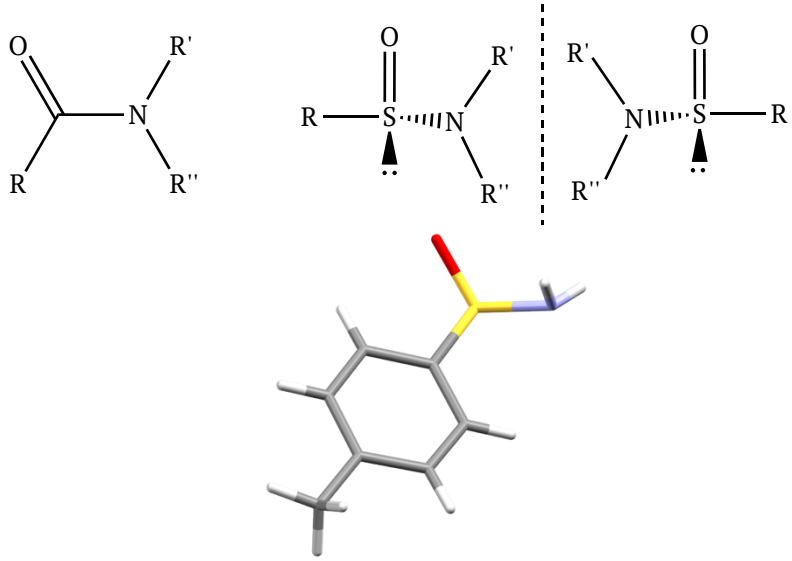


Figure 1: Comparison of the amide and sulfinamide functionalities [top] and diagram of trigonal orientation at sulfinamide nitrogen [bottom] (R'/R'' = H for primary systems).

2.1.1 Project Objectives

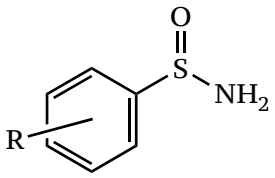
The first objective of this project was to synthesise a library of racemic primary sulfinamides, to conduct a full investigation of their solid state bonding motifs and to compare these motifs against those seen in the related series of amides in the CSD.[2] The long-term goal of this project was to explore the sulfinamides in

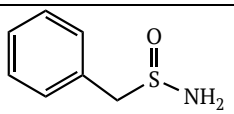
cocrystallization, but in order to conduct such a study, a clear understanding of the solid state properties of these materials must first be conducted.

2.1.2 The Synthesis of Sulfinamides

Although relatively understudied in their solid state, the sulfinamide group has received great attention in the area of asymmetric synthesis, where it is used to form enantiopure amines from sulfinimines (thio-oxime-*S*-oxides).[3] Chiral sulfinimines are accessed by condensation of enantiopure primary sulfinamides with a chosen aldehyde. Addition of organometallic reagents across the sulfinimine C-N bond produces an intermediate secondary sulfinamide which can then undergo hydrolytic cleavage to furnish a primary amine containing a new stereocentre.[3] The library of sulfinamides chosen for this study is summarised in Table 1.

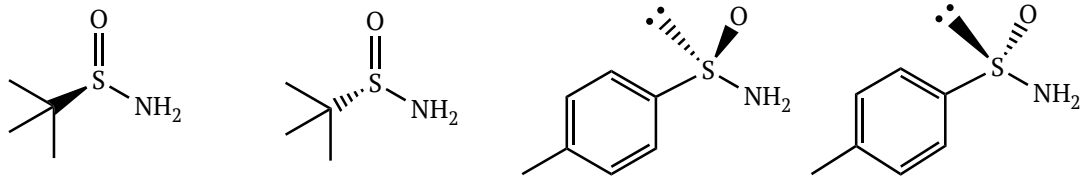
Table 1: Sulfinamide targets for this study.



Compound Number	R =
31A	H
31B[a]	
31C	2-Br
31D	3-Br
31E	4-Br
31F	2-Cl
31G	3-Cl
31H	4-Cl
31I	2-F
31J	3-F
31K	4-F
31L	2-MeO
31M	3-MeO
31N	4-MeO
31O	2-Me
31P	3-Me
31Q	4-Me

31R	2-NO_2
31S	3-NO_2
31T	4-NO_2

In asymmetric synthesis, there are two main primary sulfinamides that are used, the Ellman sulfinamides [**32**] and the Davis sulfinamides [**31Q**] (Figure 2), all of which are available commercially.

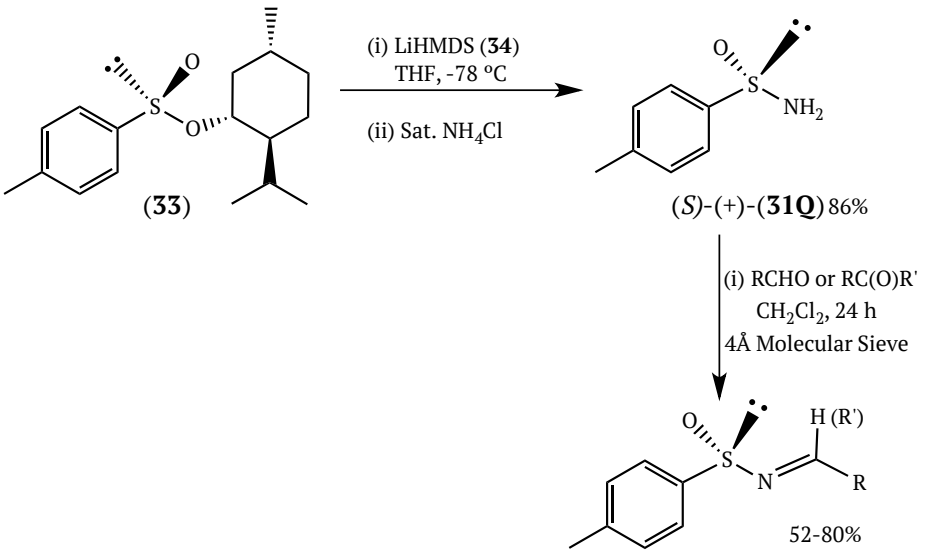


(*R*)-(+)-(**32**) (*S*)-(-)-(**32**) (*R*)-(-)-(**31Q**) (*S*)-(+)-(**31Q**)

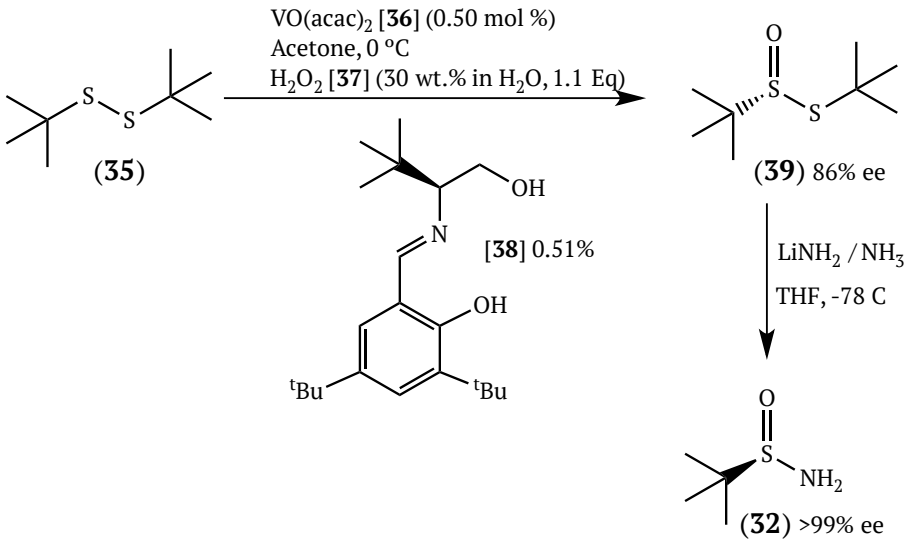
Figure 2: The Ellman [32] and Davis [31Q] sulfinamides.

In January 1997, Davis and co-workers published their synthetic strategy to form (±)-*p*-toluenesulfinamide [**31Q**] from (+) or (-)-menthyl-*p*-toluene sulfinate [**33**] as part of their synthesis of sulfinimines (Scheme 1).[3] This ester (known as the Andersen reagent) is the most widely used building block for the introduction of *p*-toluenesulfinyl groups in an asymmetric fashion.[3] The sulfinimine group can undergo a range of further reactions in the synthesis of important amine derivatives, including α-branched amines,[4] α- and β-aminophosphonic acids,[4] α,α-dibranched amines,[4] α- and β-amino acids,[4] aziridines,[4] *cis*-aziridine-2-carboxylic acids,[5] and *N*-sulfinyl aldimines and ketimines.[6]

The Ellman group published the enantioselective synthesis of (±)-*tert*-butanesulfinamide [**32**] (Ellman's sulfinamide) in June 1997 (Scheme 2), exploiting the vanadium mediated oxidation developed by Bolm and coworkers.[7,8] Ellman's sulfinamide [**32**] has been extensively used in asymmetric synthesis since this report, establishing it as a useful reagent for the general asymmetric synthesis of a broad range of amine-containing compounds (including total synthesis projects). A comprehensive review of the synthesis and applications of **32** was published in 2010.[9] Its extensive use has led to the availability of both enantiomers of **32** in large quantities at low cost from chemical suppliers.



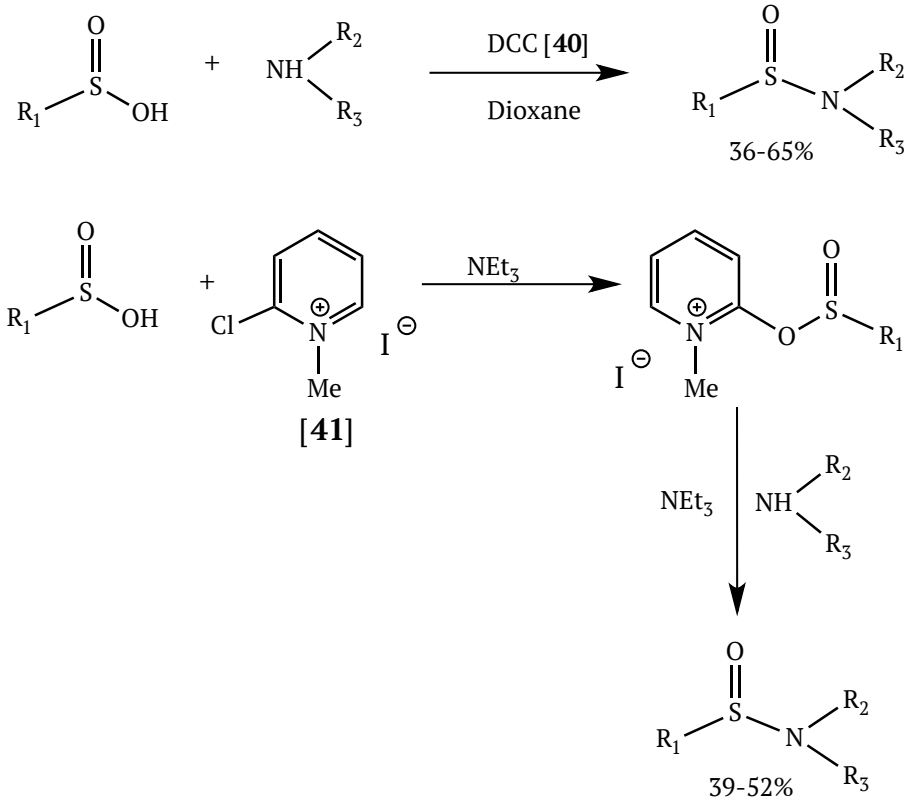
Scheme 1: The Davis route toward (*S*)-(+)-p-toluenesulfinamide [31Q], and subsequent reaction to form related sulfinimines.[3]



Scheme 2: The Ellman route to (*R*)-tert-butanesulfinamide [32] (choice of ligand in this synthesis gives the desired enantiomer).[7]

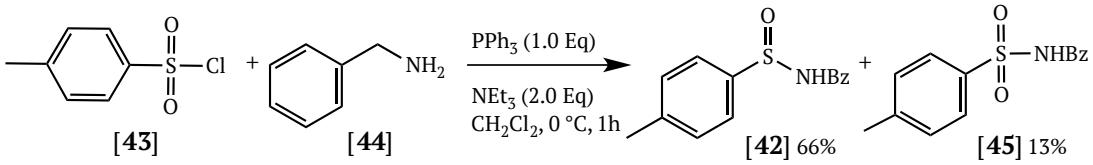
The synthesis of racemic sulfinamides has more options in the literature. Furukawa and Okawara published two synthetic methods to form secondary and tertiary sulfinamides from sulfinic acids in 1976 (Scheme 3).[10] The first route used dicyclohexyldicarbodiimide (DCC, **40**) in anhydrous dioxane at room temperature, achieving pure yields of the target sulfinamides between 36-65%. The second method used 2-chloro-1-methylpyridinium iodide [**41**] and triethylamine in dichloromethane (CH_2Cl_2) at reflux to achieve the sulfinamide in yields between 39-52%. The required series of sulfinic acids for this study was

not commercially available, and would require alternative synthetic steps to prepare, therefore this method was not chosen as the synthetic route for this project.



Scheme 3: The two synthetic strategies toward sulfinamides published by Furukawa and Okuwara (1976) (reproduced from reference).[10]

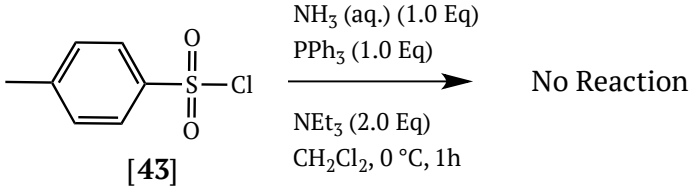
Synthesis of achiral benzyl-*p*-toluenesulfinamide [**42**] from the *p*-touenesulfonyl chloride [**43**] and benzylamine [**44**] in a one-step process was outlined by Harmata and co-workers in 2006 (Scheme 4).[11] Simple reagents (triphenylphosphine and triethylamine) were used to form the target sulfinamide [**42**] as the major product, and the corresponding sulfonamide [**45**] as the minor product. These products could be easily separated using column chromatography, and this method had been previously used with success within our research group.[12]



[**43**] [**44**] [**42**] 66% [**45**] 13%

Scheme 4 'One-pot' method of (±)-p-toluene sulfinamide [31Q] preparation (reproduced from Harmata et al.). [11]

During our work, it was initially postulated that if aqueous ammonia could be used as the amine source, the Harmata strategy could be utilised to prepare the target sulfinamides from commercially available sulfonyl chlorides. This reaction was used as a first attempt at one-pot primary sulfinamide synthesis using *p*-toluenesulfonyl chloride [**43**] and aqueous ammonia solution as the amine source in place of benzylamine [**44**] (Scheme 5).



[**43**]

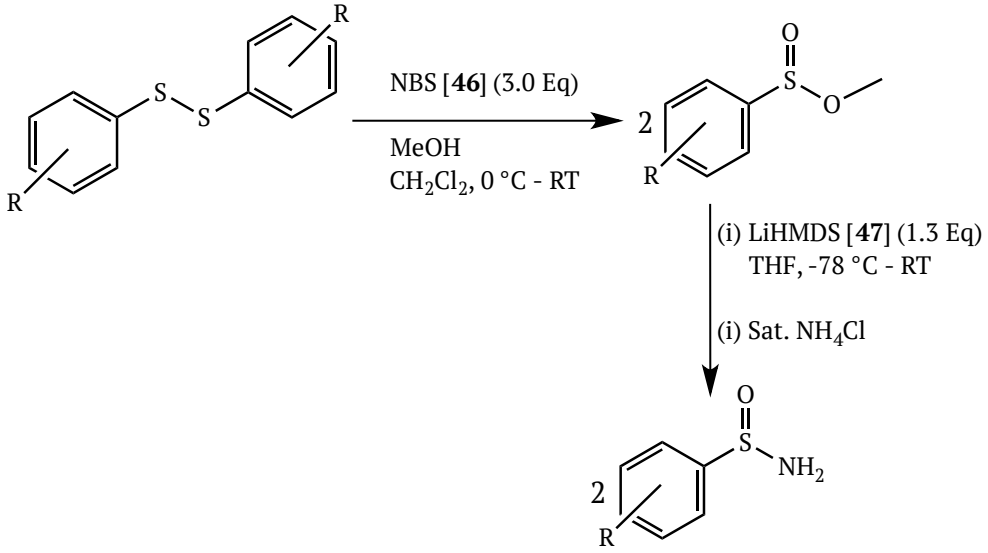
Scheme 5: Attempted synthesis of *p*-toluene sulfinamide [31Q].

The solvent used in the literature was CH_2Cl_2, and so the reaction mixture was a biphasic system of **43** in CH_2Cl_2, and the aqueous ammonia solution. The reaction mixture was stirred vigorously, however, no evidence of product formation was observed in this case, with only the starting sulfonyl chloride recovered from solution.

If required, the Ellman and Davis methods of synthesis can be adapted to form racemic primary sulfinamides in high yields. The chiral ligands and vanadium catalyst used in Ellman's oxidation method can be bypassed in favour of achiral hydrogen peroxide,[9] and the chiral menthyl sulfinates substituted by achiral sulfinate esters in the Davis methodology.[3,6] These methods would potentially yield the target primary sulfinamide in two steps. An alternative route to sulfinamides is treatment of an amine with a suitable base followed by nucleophilic displacement of the menthyl moiety from the Andersen reagent[13] [**33**] or a similar sulfinate ester, however, this route was not chosen due to the requirement for strongly basic conditions in order to deprotonate ammonia,

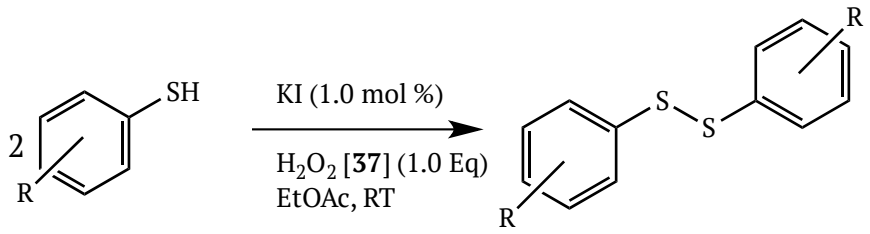
coupled with the availability of more attractive synthetic routes (discussed below).

In 2004, Garcia Ruano *et al.* published a suitable synthetic strategy to form racemic primary sulfinamides from disulfides on a multi-gram scale in high yields (Scheme 6).[14] The strategy utilised the synthesis of an achiral methyl sulfinate ester using *N*-bromosuccinimide (NBS) [**46**] as performed by Brownbridge and Jowett,[15] followed by conversion of the ester to the target primary sulfinamide using lithium bis(hexamethyldisilyl)amide (LiHMDS) [**47**] using the procedure proposed by Davis (Scheme 1). We decided to utilise this straightforward approach to access the series of sulfinamides for this study.



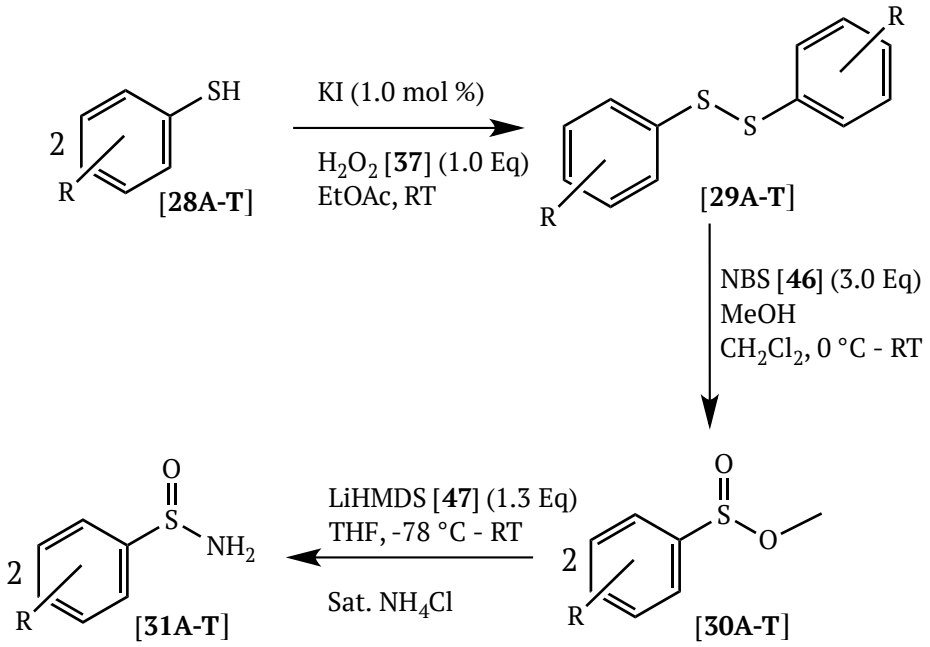
Scheme 6: **Synthetic route to sulfinamides published by Garcia Ruano *et al.*[14]**

The complete series of substituted disulfides was not available commercially. However, it was envisaged that the remaining substituted aryl-disulfides could be accessed easily *via* an iodide-catalysed thiol oxidation using hydrogen peroxide, following the procedure published by Kirihara *et al.* in 2006 (Scheme 7).[16]



Scheme 7: Synthetic route to disulfides (Kirihara et al.).[16]

The combined synthetic strategy envisaged for this project would form the target sulfinamides over two or three steps depending on the commercial availability of the parent disulfides (Scheme 8).



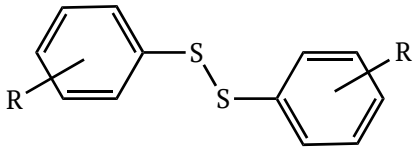
Scheme 8: **The combined synthetic strategy to access the library of sulfinamides [31A-T].**[14–16]

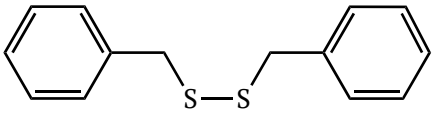
2.2 Synthesis

2.2.1 Synthesis of Aryl Disulfides [29A-T]

The synthetic method outlined by Kirihara *et al*. (Scheme 7) was used successfully in all cases for this work. Table 3 illustrates the library of disulfides synthesised in this study. Compounds **29A**, **29B**, **29N**, **29R**, **29S**, and **29T** (Table 2) were purchased from chemical suppliers.

Table 2: Commercially available disulfides used in this work

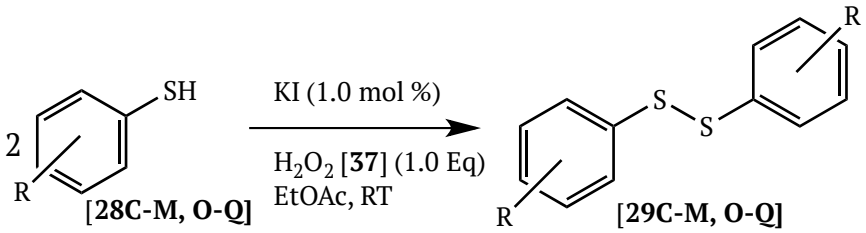




Compound Number	R-Group
29A	H
29N	4-MeO
29R	2-NO_2
29S	3-NO_2
29T	4-NO_2

29B

Table 3: Percentage yields of disulfides [29C-M, O-Q] prepared in this work.



Compound Number	R - Group	Yield	Appearance
29C[17]	2-Br	88%	White Solid
29D	3-Br	91%	Orange Oil
29E[18]	4-Br	84%	Yellow Solid
29F[17]	2-Cl	90%	Off-white Solid
29G[19]	3-Cl	87%	Orange Oil
29H[19]	4-Cl	92%	Yellow Solid
29I[20]	2-F	91%	Yellow Oil
29J	3-F	99%	Yellow Oil
29K[20]	4-F	85%	Yellow Oil
29L[21]	2-MeO	91%	Off-white Solid
29M[21]	3-MeO	96%	Red-brown Oil
29O[22]	2-Me	89%	Yellow Solid

29P[23]	3-Me	90%	Yellow Oil
29Q[16]	4-Me	72%	Yellow Solid

While the Kirihara method utilised sodium iodide with hydrogen peroxide for the catalytic oxidation to the disulfides, potassium iodide was used in its place for this work without issue. Other than this substitution, the literature method was followed exactly. The products were obtained in very good to near-quantitative yields on a 1-5 g scale; after workup with aqueous sodium thiosulfate to remove excess iodine. In all instances the use of these disulfides proved successful and so chromatographic purification was not explored. These disulfides proved very stable with respect to storage at room temperature for a number of months.

Of the 14 disulfides prepared, two were novel compounds (**29D** and **29J**) and were fully characterised. Characterisation of the non-novel materials was performed using ^{1}H, ^{13}C NMR, and IR, as standard. The aromatic splitting pattern observed in ^{1}H NMR was reflective of the 1, 2 or 3 substitution on the aromatic rings in each case. Infrared spectroscopy, despite its use as a standard characterisation technique, provides little diagnostic information for disulfide materials; the disulfide *S-S* stretch is weak and appears in the fingerprint region of the spectrum (500-700 cm^{-1}).[24] There are few other characteristic stretches in the spectra, and so nominal mass spectrometry was performed on all of the materials as a supplementary analytical technique. The disulfides fragmented to a great degree under positive mode mass spectrometry conditions; the parent molecular ion [$(M+H)^{+}$] was not detected in almost all cases. The negative electron-spray ionisation mode, however, provided valuable information; fragmentation did occur, but the monomer (R-$C_6H_6S^{-}$) could be detected for all materials, and was used as an indicator of the parent disulfide compound in the absence of [$(M+H)^{+}$].

In the case of **29H, 29I**, **29K**, and **29M**, the literature reported these materials as white solids. **29I** and **29K** were both yellow oils in our hands, and slight differences were recorded in the NMR spectra when compared to that reported in the literature.[20] Microanalysis was successfully performed on these materials to confirm the presence of the desired products. In the cases of **29H** and **29M**, the materials were, again, reported as a white solids in the literature (found to be

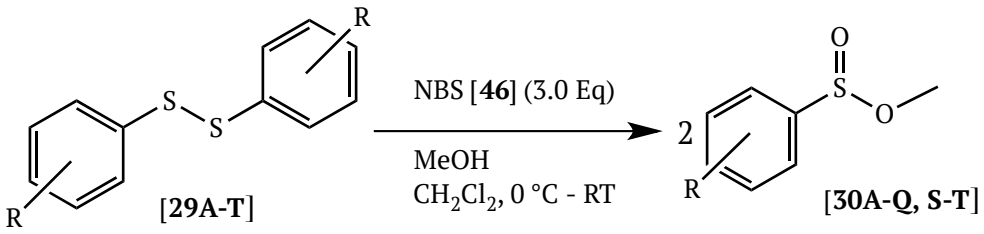
a yellow soid and a red-brown oil respectively).[19,20] However, the spectral characteristics recorded for these agreed well with those reported and so microanalysis was not performed in these cases. **29H** is reported as a white solid, and was found to be a yellow solid in this project, it is possible that a small amount of residual iodide or iodine remaining in the material could account for this colour.

2.2.2 Synthesis of Methyl Aryl Sulfinate Esters [30A-Q, S--T]

The synthetic method outlined by Brownbridge and Jowett[15] was used for all 20 disulfides [**29A-T**] (Table 4) including the 6 commercially available disulfides. The method used methanol, CH_2Cl_2 and NBS [**46**] (which was freshly re-crystallized from hot water before use).[25] The sulfinate esters were prepared in moderate to good yields on a 1-5 g scale, and appeared to suffer structural degradation with storage at room temperature. Upon discovering the structural lability of these compounds, it became customary to utilise the material within 24 h of preparation.

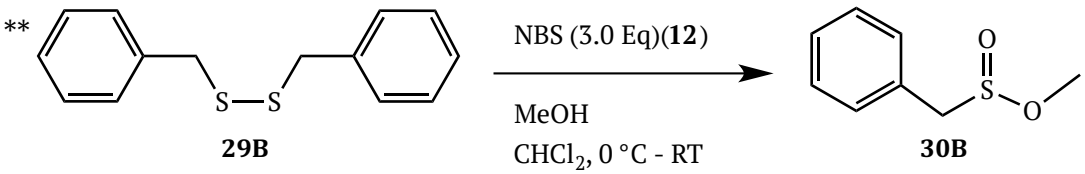
This 'one-pot' method[15] proved very effective for preparation of the target methyl sulfinate esters, of the 20 target esters attempted, 19 were successfully prepared. Column chromatography was required in 14 of the 19 syntheses. 5 materials (**30A**, **30B**, **30D**, **30E**, and **30J**) did not require purification, as they appeared clean in ^{1}H and ^{13}C NMR after reaction work-up with aqueous sodium bicarbonate and water. Chromatography of these materials was time-consuming, owing to the fact that the materials tended to streak on the silica gel and were isolated from the column over a large volume of solvent (500 mL to 1 litre). Many alternative solvent systems were investigated *via* TLC, but this issue could not be resolved. Despite this small challenge, the chromatography method proved very effective in purifying the materials to an analytically pure state.

Table 4: Percentage yields of methyl aryl sulfinate esters [30A-Q, S-T] (those derived from commercially available disulfides in green).



Compound	R=	Percentage Yield	Appearance
30A[a,26]	H	61%	Colourless Oil
30B[a,15]	**	85%	Colourless Oil
30C[b]	2-Br	68%	Yellow Oil
30D[a]	3-Br	71%	Orange Oil
30E[a,27]	4-Br	73%	Colourless Oil
30F[b,28]	2-Cl	81%	Colourless Oil
30G[b,27]	3-Cl	56%	Colourless Oil
30H[b,29]	4-Cl	89%	Yellow Oil
30I[b]	2-F	59%	Colourless Oil
30J[a]	3-F	89%	Colourless Oil
30K[b,26]	4-F	50%	Yellow Oil
30L[b,14]	2-MeO	77%	Pale yellow Oil
30M[b]	3-MeO	90%	Pale yellow Oil
30N[b,27]	4-MeO	74%	Pale yellow Oil
30O[b,28]	2-Me	71%	Pale yellow Oil
30P[b]	3-Me	74%	Yellow Oil
30Q[b,5,30]	4-Me	83%	Colourless Oil
30R[15]	*2-NO_2*	*No Reaction*	-
30S[b,31]	3-NO_2	77%	Orange Oil
30T[b,14]	4-NO_2	61%	Orange Solid

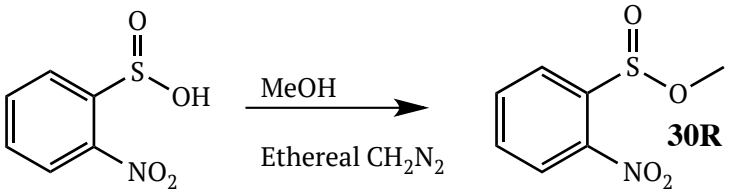
a - Additional purification not required; b - Column chromatography on silica gel performed.



In all successful cases, the addition of NBS [**46**], resulted in a gradual evolution of a strong orange colour over 1-5 minutes, due to liberated bromine in the reaction. This colour was used as an indicator of reaction initiation and progress, as it

tended to disappear as the reaction progressed toward completion. TLC monitoring was used to determine reaction completion, with disappearance of the disulfide band used for reference. The orange colour was not observed in the case of 2-nitrophenyl disulfide [**29R**] and disappearance of the disulfide band was not seen in TLC. Only starting material was recovered from the reaction vessel. The solubility of the 2-nitrophenyl disulfide [**29R**] was an issue in this case; the normal 1:1 methanol:CH_2Cl_2 solvent system was used here, but additional CH_2Cl_2 was required to fully dissolve **29R**. Tetrahydrofuran (THF) was used as an alternative solvent for the synthesis of the sulfinate target, but without success. Although the **29R** dissolved, no product was formed in the reaction. The literature was consulted again, and synthesis of **30R** was attempted using sodium carbonate and liquid bromine in methanol as described for the preparation of methyl benzene sulfinate [**48**] by Meyers and Resek,[32] however, no product was formed using this method.

The crystal structure of methyl 2-nitrobenzene sulfinate [**30R**] is available in the CSD,[2] published by Kucsman and Kapovits in 1989.[33] This methodology used an ethereal solution of diazomethane in methanol to generate **30R** (Scheme 9) which was subsequently characterised by X-ray crystallography.



Scheme 9: Synthetic route to methyl 2-nitrobenzene sulfinate [30R] [adapted from Kucsman and Kapovits].[33]

Since diazomethane is classified as a high risk material,[34] it was decided that continuation of the library of 19 sulfinate esters in the absence of methyl 2-nitrobenzene sulfinate [**30R**] was the best way to move forward with the project at this point.

The sulfinate esters prepared were easily characterised by ^{1}H, ^{13}C NMR and IR; nominal and high resolution mass spectrometry was performed on materials that required full characteristation. The sulfinate ester moiety resulted in two characteristic stretches in the IR spectra, the S=O stretch appeared between

1112-1183 cm^{-1} and the S-O between 957-999 cm^{-1}. The introduction of the methyl ester in the compounds resulted in a characteristic 3H singlet at approximately 3.50 ppm (3.48-3.58 ppm) in the ^{1}H NMR spectra, and a new CH_3 signal appearing at ~50.0 ppm (49.8-56.8 ppm) in the ^{13}C NMR spectra.

Upon conversion to the sulfinate ester, quaternary carbon signals were identified and labelled based on the large downfield shift of the quaternary carbon bearing the sulfur to ~140 ppm (128.8-150.5 ppm). It is sensible to conclude that the carbon bearing the sulfur would represent the larger quaternary carbon shift in the spectrum, having sustained a considerable molecular change at that position. This shift lead to re-assessment of all carbon signal assignments for the disulfides [**29C-M, O-Q**]. For all sulfinates [**30A-Q, S-T**] synthesised in this study, the carbon (Cq) signals were retrospectively compared to their disulfide precursors to ensure correct assignment of the quaternary carbon signals for the disulfides and sulfinate esters.

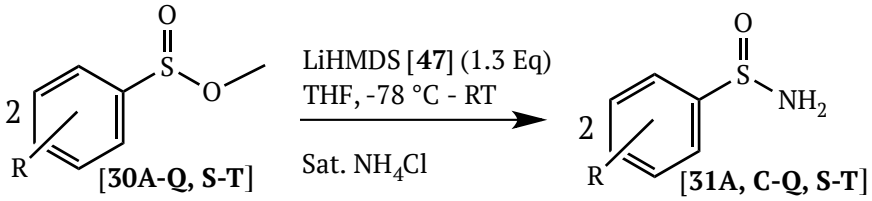
Of the 19 sulfinate esters prepared, 9 were not found in the literature at the time of preparation (**30C**, **30D**, **30E**, **30G**, **30I**, **30J**, **30M**, **30N**, and **30P**) and so these materials were fully characterised. A study was published in early 2016 that contained some characterisation information for three sulfinate esters contained in this study (**30E**, **30G**, and **30N**).[27] In the case of **30D**, the ^{1}H NMR spectrum for the material was reported as part of a patent,[35] however, no other characterisation was reported and so full characterisation of this material was performed. There was a slight shift noticed between the literature and recorded NMR spectra for **30L** (0.16 ppm in the ^{1}H NMR, and 0.7 ppm in the ^{13}C NMR). However, this was attributed to a difference in automatic calibration of the reference peak by NMR software. Nevertheless, full characterisation was performed on this material, with high resolution mass spectrometry confirming the presence of the desired structure. A difference was also noticed between the literature (recorded in 1978)[31] and our spectrum for **30S**. There was a much higher level of peak resolution determined for our material, again with full characterisation performed to confirm the presence of the target compound.

2.2.3 Synthesis of Aryl Primary Sulfinamides [31A-T]

The final step in the synthesis of the target materials was conversion of the methyl sulfinate esters to their corresponding primary sulfinamides according to the Davis procedure.[36] The method described use of a solution of the Andersen reagent [**33**] in dry THF at -78 °C, followed by the addition of LiHMDS [**47**] before warming the reaction to room temperature. In order to form target materials [**31A-Q, S-T**], LiHMDS [**47**] was added to a cooled solution of the racemic methyl sulfinate ester [**30A-Q, S-T**] in dry THF.

The Davis method was used to great success, synthesising 18 of the 19 sulfinamides. The sulfinamide products were successfully synthesised on 0.5-2 g scale, in moderate to good yields. Purification was performed, where required, by either column chromatography, recrystallization and/or ammonium chloride washes, providing pure products in all cases (Table 5). Chromatography presented similar challenges to the sulfinate esters discussed above, products tended to streak from the column, regardless of solvent system, and were isolated over large volumes (approx. 500-700 mL). Of the 18 sulfinamides prepared, seven were not previously characterised in the literature (**31C**, **31D**, **31G**, **31I**, **31J**, **31M**, and **31P**) and were fully characterised. A range of NMR solvents are reported in the literature for characterisation of the sulfinamides, leading to slight differences in the reported NMR spectra for **31E**, **31N**, **31Q**, and **31S**; the observed spectra for these materials were consistent with the introduction of the NH_2 group and the relative substitution of the aromatic rings in each case. Full characterisation of the materials was performed as a matter of caution, and the presence of the desired material was confirmed in all cases.

Table 5: The percentage yields for sulfinamides [31A-T] prepared in this work (novel materials shown in green).



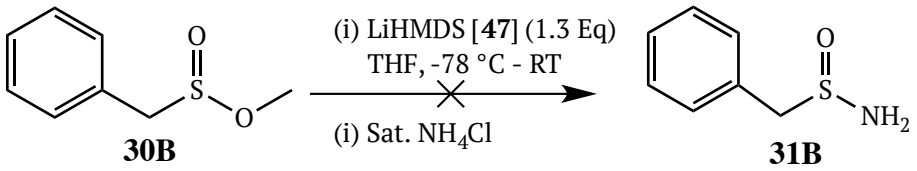
Sulfinamide Product	R =	Percentage Yield	Appearance
31A[b,37]	H	72%	Off-white Solid
31B	**	*No Reaction*	-
31C[b]	2-Br	19%	Off-white Solid
31D[a]	3-Br	87%	Yellow Solid
31E[b,38]	4-Br	50%	Off-white Solid
31F[a,28]	2-Cl	87%	Cream Solid
31G[a]	3-Cl	94%	Cream Solid
31H[a,37]	4-Cl	80%	Yellow-Cream Solid
31I[a]	2-F	98%	Off-white Solid
31J[a]	3-F	91%	Orange-yellow Solid
31K[d,28]	4-F	82%	Yellow-Cream Solid
31L[b,c,14]	2-MeO	72%	White Solid
31M[b]	3-MeO	67%	Orange Solid
31N[d,37]	4-MeO	87%	Cream Solid
31O[b,28]	2-Me	31%	Cream Solid
31P[a]	3-Me	93%	Cream Solid
31Q[a,29,37]	4-Me	81%	White Solid
31S[a,31]	3-NO_2	75%	Orange Solid
31T[b,14]	4-NO_2	72%	Orange Solid

a - Additional purification not required; b - Column chromatography on silica gel performed;
c - Recrystallization performed; d - Additional NH_4Cl wash performed.
** See Scheme 10.

Structural confirmation of the sulfinamide materials was completed using the normal suite of methods (^1H and ^{13}C NMR, IR, nominal/high resolution mass spectrometry). The formation of the sulfinamide moiety was most easily detected in the ^1H NMR, with the appearance of a broad 2H singlet at ~6.0 ppm (5.0-6.57 ppm). Two new N-H stretches were observed in the IR spectra for these materials at approximately 3300 and 3150 cm^{-1}. The change in the ^{13}C signal for the quaternary carbon bearing the sulfur was only a slight shift in some of the

cases, but comparison between the parent disulfide, sulfinate ester and sulfinamide ^{13}C NMR spectra allowed for correct interpretation of the signals in all cases.

In the case of phenylmethane sulfinamide, the desired product could not be formed (Scheme 1); most likely due to deprotonation of the benzyl protons by the **47**. In an attempt to circumvent this issue, the reaction was repeated using an increased loading of **47** (2.2 equivalents as opposed to the initial 1.3 equivalents). This method, however, was also unsuccessful, generating an unidentifiable mixture of materials with no evidence of formation of the sulfinamide product [**31B**].



30B **31B**

Scheme 1: Attempted synthesis of methyl phenylmethane sulfinamide (31B).

With this in mind, the structural stability of the methyl phenylmethane sulfinate sample was investigated using ^{1}H NMR, and it was observed that the structural integrity of the sulfinate ester (**30B**) degraded quickly with storage at room temperature (Figure 3). The degradation in the NMR presents a characteristic peak at approximately 10 ppm, consistent with the formation of benzaldehyde. The formation of benzaldehyde was further corroborated by a slight almond odour from the sample retain with storage over time. The only other synthesis of phenylmethane sulfinamide contained in the literature utilises benzyl mercaptan and chloramine at 0 °C as described by Erlenmeyer and Seiler in 1957.[40] This synthetic route was not considered due to the safety concerns surrounding the use of chloramine. The only other published synthesis of phenylmethane sulfinamide is an enantiospecific route published as part of a Chinese patent in 2001, which was not considered due to the requirement for a racemic synthesis and the need for translation of the entire document;[41] and so synthesis of the phenylmethane sulfinamide was not pursued further.

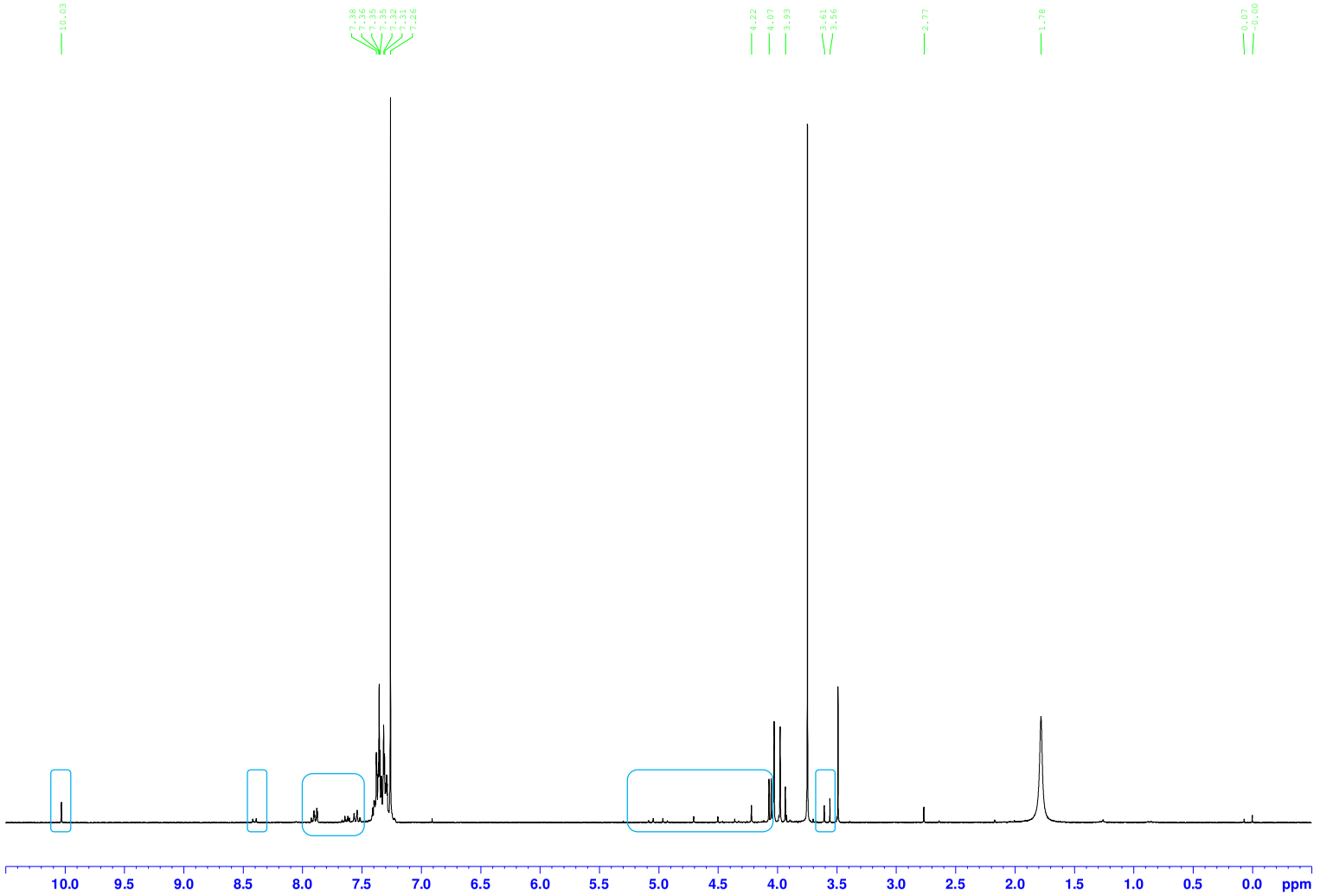


Figure 3: ^{1}H NMR of methyl phenylmethane sulfinate (30B) after one week of storage at room temperature. Signs of structural change are highlighted in blue.

2.2.4 Conclusions

The synthetic route to primary sulfinamides was used effectively to prepare 14 disulfides, of which 2 were novel; 19 sulfinate esters, of which 8 were novel; and 18 primary sulfinamides, of which 7 were novel. The target sulfinamide materials were prepared in 2 or 3 synthetic steps in good yields and purity and, if stored as solids, were bench-stable for several months; with no evidence of structural degradation (unlike the labile sulfinate intermediates). 2 of the 20 target materials could not be accessed using our chosen methods, and although alternative synthetic options were attempted, successful preparation of phenylmethane sulfinamide and 2-nitrobenzene sulfinamide could not be completed.

Simple purification methods such as column chromatography were used to great effect in this project, and characterisation of these materials was performed via IR, ^{1}H NMR, ^{13}C NMR. Nominal mass spectrometry and high resolution mass spectrometry or microanalysis was performed for novel materials.

2.3 Structural Characterisation of Primary Sulfinamides

2.3.1 Sulfinamides and Amides in the Solid State

As mentioned in section 2.1, sulfinamides have received scant attention in relation to their solid state properties. This leaves an important gap to be filled in our current understanding of solid state motifs since sulfinamides represent a chiral relative of the amide functionality. A search of the CSD[2] reveals only four known structures of primary sulfinamides (each will be discussed in section 2.3.3). When a comparable search of the CSD[2] is performed for the primary amide moiety 2546 structures are returned.

Using hydrogen bond donor (α) and acceptor (β) constants from Hunter's table[41] as a numerical guideline, it can be seen that the sulfinamide is both a potent hydrogen bond donor (N-H, α = 3.2) and hydrogen bond acceptor (S=O, α = 8.3) (Table. 6). Acids and amides are common functional group targets in crystal engineering studies.[42–44] Therefore, it is informative to compare these donor and acceptor constants to the sulfinamide functional group (Table 6).

Table 6: Hydrogen bond donor (α) and acceptor (β) constants for sulfinamide, carboxylic acids and amides (adapted from reference)[41]

Functional group	H-Bond donor (α)	H-Bond acceptor (β)
Sulfinamide	3.2	8.3
Carboxylic Acid	3.6	5.3
Amide	2.9	8.3

For hydrogen bond donor value (α), sulfinamides reside between the carboxylic acid (O-H, α = 3.6) and the amide functional groups (N-H, α = 2.9). The sulfinamide (S=O) acceptor value (β) is identical to that of the amide (S=O/C=O, β = 8.3), and lies far above that of the carboxylic acid (C=O, β = 5.3). This suggests that the sulfinamide functional group should form strong N-H···O=S hydrogen bond interactions in the solid state, further reinforcing the need for investigation of the solid state behaviour of these materials.

2.3.2 Primary Amides in the CSD[2]

The intended approach of this work is to compare the crystal structures obtained for primary aryl sulfinamides with their amide counterparts in the CSD.[2] The CSD[2] contains 2,507 crystal structures containing the primary amide moiety, of which 409 are primary aryl amides of the general structure shown in Figure 4.

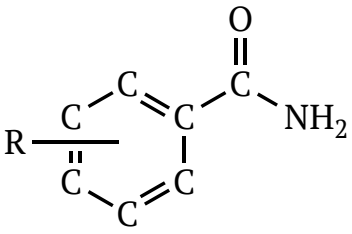


Figure 4: General format for primary amides used in CSD[2] searches, (all benzene ring substitutions allowed, no aryl-hydrogens defined).

Hydrogen bonding in amides occurs through the strong C=O···H-N interaction between the carbonyl oxygen and the amide hydrogens. If this primary aryl amide sub-set is inspected for recurrent bonding motifs using Mercury,[45] two main motifs can be identified. The $R_2^2(8)$ dimer is a particularly strong and recurrent bonding motif, seen in 216 of the 409 structures [Figure 5 (left)]. The other dominant bonding motif is the C(4) chain [Figure 5 (right)].

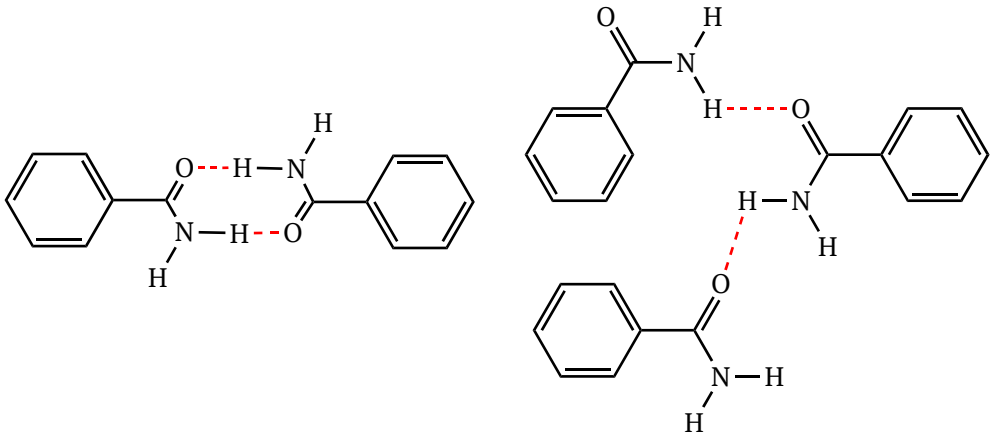


Figure 5: The $R_2^2(8)$ dimer (left) and C(4) chain (right) in primary aromatic amides (hydrogen bonds shown in red).

Analysis of packing modes in primary amides as carried out by Leiserowitz and Schmidt remarked that the centrosymmetric $R_2^2(8)$ motif, consisting of two donors and acceptors is the 'ideal' situation for close packing in the solid state of primary amides, with the majority forming these pairs, having the centre of the

dimer coinciding with centres of symmetry in the unit cell.[46] The alkyl/aryl-group attached to the primary amide moiety plays a significant role in allowing this centrosymmetric relationship. For example, enantiopure primary amides would not be able to crystallize across a centre of symmetry in this way, and the probability of finding a symmetrical dimer-type structure in these molecules is low due to the chiral alkyl/aryl group.[47]

The N-H···O=C interaction tends to adopt a motif that is planar, involving two hydrogen bonds from N-H···O=C.[48] A good example of the robustness of this motif is that of benzamide [**13**], which displays the $R_2^2(8)$ dimer across an inversion centre (Figure 6) in all 3 polymorphic forms, with differences in these forms occurring due to π-π interactions.[49]

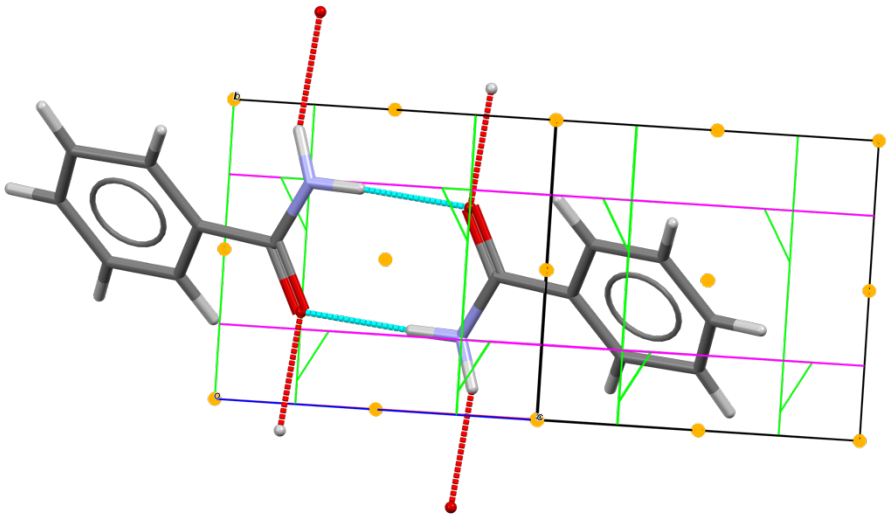


Figure 6: $R_2^2(8)$ dimer motif in 13 Form I (BZAMID01), symmetry elements shown – inversion centres (yellow), 2_1-screw axis (green), glide plane (magenta).[50]

The second prevalent bonding motif is the C(4) chain, occurring frequently in this subset of primary amide structures, and again seen in benzamide [**13**] (Figure 7). Chain motifs are an attractive target in crystal engineering since they allow controlled assembly in one dimension, adding directionality to the crystal growth.[51] Studies have shown that, depending on the specific availability of other hydrogen bond acceptors, either the $R_2^2(8)$ dimer or the C(4) chain motifs can be broken.[52,53]

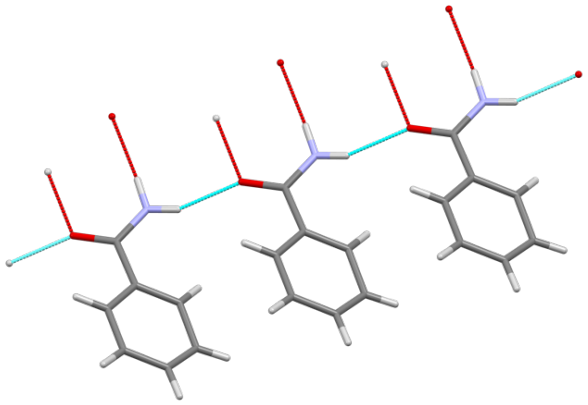


Figure 7: C(4) chain motif in 13 form I (BZAMID01).[50]

When the $R_2^2(8)$ dimer and C(4) chain motifs combine in the solid state, another ring is formed, an $R_4^2(8)$ tetramer; this combination of motifs is frequently referred to as the "amide ladder".[52] This ladder motif can be seen repeatedly throughout the literature and is a characteristic feature of amide crystal packing (Figure 8).[52]

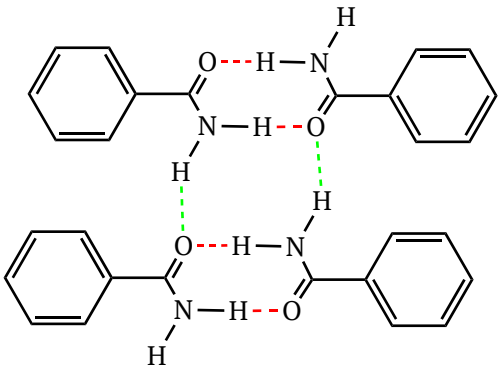


Figure 8: The binary level 'amide ladder' motif; $R_2^2(8)$ dimer bonds in red, C(4) chain bonds from the *anti* N-H in green. $R_4^2(8)$ tetramer formed in the centre of the image.

If suitable alternative hydrogen bonding paths are presented, both the dimer and the chain motifs ordinarily displayed by amides can be disrupted, for example when NO_2, CN, Cl, Br, or OMe groups are present within the structure.[52] This results in entirely different packing motifs. 4-Nitrophthalamide [**49**], for example, displays the anticipated $R_2^2(8)$ dimer interaction, but the C(4) chain has been broken, in favour of N-H···O=N interactions (Figure 9).

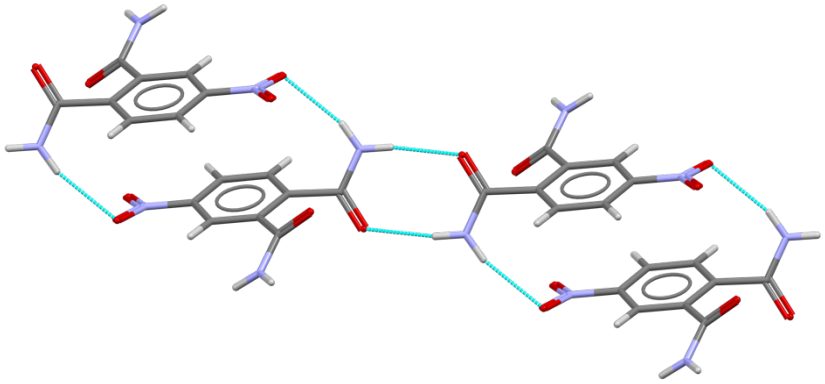


Figure 9: The disruption of the C(4) chain motif in 4-nitrophthalamide [49] (VIZZIP).[54]

In the case of 2,4,6-trinitrobenzamide [**50**] the molecules participate in discrete C-H···O=N and N-H···O=N interactions to form the crystal structure and the anticipated dimer motifs are not observed at all (Figure 10),[55] instead a C(4) chain is seen, utilizing one of the two N-H bonds and a more complex C(7) chain is seen through N-H···O=N interactions. These C(7) chains combine at a binary level to form a large $R_4^4(18)$ tetramer.

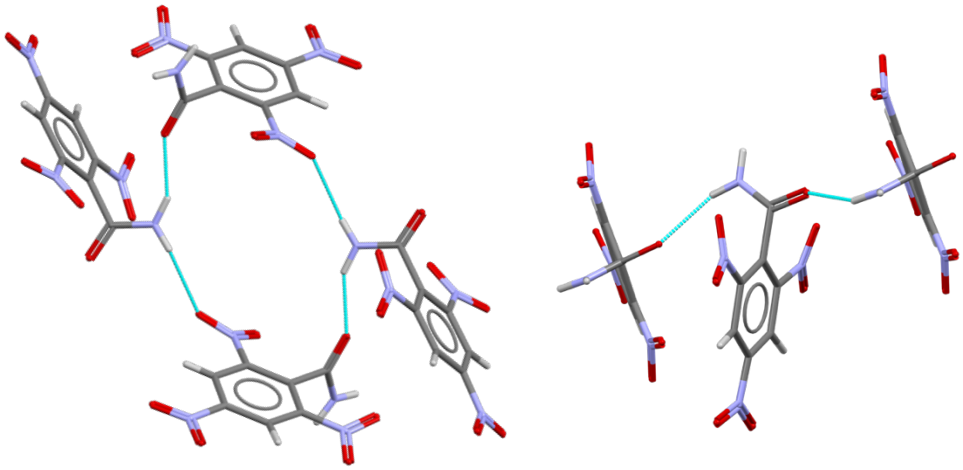
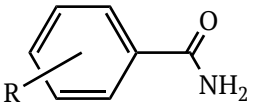
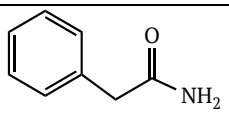


Figure 10: Complete disruption of the amide ladder in 2,4,6-trinitrobenzamide[50] forming C(4) chains (right) and a binary level $R_4^4(18)$ tetramer (left) (SAWBOH).[55]

In summary, it can be concluded from the literature that the $R_2^2(8)$ dimer and C(4) chain are the primary motifs of interest, in the context of primary aromatic amide crystallographic packing. However, when presented with alternative hydrogen bonding pathways, these robust synthons can become less dependable.

To build upon this information, the structural data for the direct amide-counterparts [**51A-T**] of the substituted sulfinamides chosen for this work were examined in detail, and their interactions investigated for similarities/differences to that shown above. Table 7 shows the distribution of the two main binary interactions among this sub-set of primary aryl amides, along with the occurrence of the amide ladder. The set of chosen amide structures contained many polymorphic structures, but only those polymorphs for which differences in the hydrogen bonding motifs have identified are included in Table 7.

Table 7: The distribution of hydrogen bonding motifs among select primary aryl amides in the CSD.[2]

Compound Number		$R_2^2(8)$ dimer present	C(4) chain present	Ladder present	CSD Refcode[2]
51A[a]	**H**	✔	✔	✔	BZAMID01[49]
51B		-	-	-	Not in CSD
51C	**2-Br**	✔	✔	✔	BRBZAO[56]
51D	**3-Br**	**X**	✔	**X**	BRBZAM[57]
51E[a]	**4-Br**	✔	✔	✔	BBEZAM02[58]
51F[a]	**2-Cl**	✔	✔	✔	CLBZAM10[59]
51G	**3-Cl**	✔	✔	✔	NABRAJ[60]
51H[a]	**4-Cl**	✔ ✔	✔ ✔	**X** ✔	PCBZAM03[61] PCBZAM01[62]
51I	**2-F**	✔	✔	✔	BIGSUF[63]
51J[a]	**3-F**	✔	✔	✔	BENAFM10[64]
51K[a]	**4-F**	✔ ✔	✔ ✔	**X** ✔	BENAFP01[65] BENAFP02[66]
51L[a]	**2-MeO**	✔	✔	✔	RECQIA[67]
51M	**3-MeO**	-	-	-	Not in CSD
51N	**4-MeO**	-	-	-	Not in CSD
51O	**2-Me**	✔	✔	✔	NABQEM[68]
51P	**3-Me**	✔	✔	✔	MEBENA[69]
51Q[a]	**4-Me**	✔	✔	**X**	DABVAD01[70]
51R	**2-NO_2**	✔	✔	✔	ONBZAM[71]
51S	**3-NO_2**	✔	✔	**X**	JACYOB[72]
51T[a]	**4-NO_2**	✔	**X**	**X**	NTBZAM01[73]

a – Alternative polymorphic forms of this material exist, however the hydrogen bonding motifs do not differ from the form(s) listed in the table;

Clearly, the $R_2^2(8)$ dimer and C(4) chains are very robust structure determining features, with these motifs observed together in all but 2 of the 17 structures (the remaining 2 structures containing one motif each). The amide ladder is observed in 13 of these examples, with 4 displaying combinations of the dimer and chain that do not form a ladder. It is interesting to note that in the cases of the polymorphic structures **51H** (2-Cl) and **51K** (4-F), the $R_2^2(8)$ dimer and C(4) chain persist in all polymorphs, and the polymorphic variation arises only from the presence or absence of the ladder. This highlights the relative persistence of the simpler unitary features compared to the more complex binary amide ladder.

The structure of BRBZAM[57] is not fully complete, in that the hydrogens are not present and the R-factor is 11%, therefore it is difficult to confirm that the amide ladder is not present in this case, although it appears that the crystal structure is supported by linear N···O=C bonds and weak C-H···π interactions. The disruption of the ladder motif in the cases of JACYOB[72] [**51S**] and NTBZAM01[73] [**51T**] seems to be due to preferential interaction of the amide N-H with the nitro group, similar to that observed for 4-nitrophthalamide discussed previously (Figure 9).

The 4-methyl analogue, DABVAD01[70] [**51Q**] displays both the C(4) chain and the $R_2^2(8)$ dimer, but these chains do not form the amide ladder. The $R_2^2(8)$ dimer is formed between the amide moieties with only one of the two available hydrogens participating in the dimer, the other hydrogen atom participates in the C(4) chain; creating a 4-point interaction for each dimer *via* C(4) chains (Figure 11). This 4-point interaction can also be observed in BENAFP01[65] [**51K**] and PCBZAM03 [**51H**].[61]

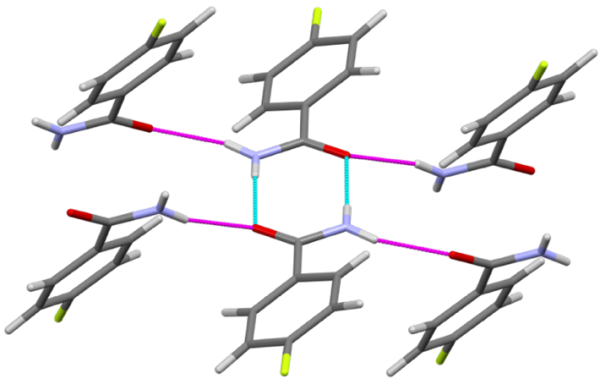


Figure 11: 4-point interaction in amides [C(4) chains in magenta and $R_2^2(8)$ in cyan].

In the context of sulfinamide crystallographic packing, it is unlikely that the planar $R_2^2(8)$ dimer motif will be seen, since the chirality of the sulfinamide functional group, combined with the trigonal orientation around the nitrogen, will not allow the required planarity of the interaction. However, the more easily broken C(4) chain motif is expected to be observed.

2.3.3 Sulfinamides in the CSD[2]

The S=O···H-N hydrogen bond is strongly apparent in the literature, with 4284 examples contained in the CSD.[2] As mentioned in section 2.3.1, the CSD[2] contains only four known primary sulfinamide crystal structures. Two of these structures, JAZZIT[74] [**52**] and VICGET[75] [**53**] (Figure 12) are large ribonucleoside derivatives, and the other two, XEPHOR[76] [**54**] and ZIJGOQ[1] [**31Q**] are small-molecule primary sulfinamides; the latter is a member of the set of sulfinamide targets for this study.

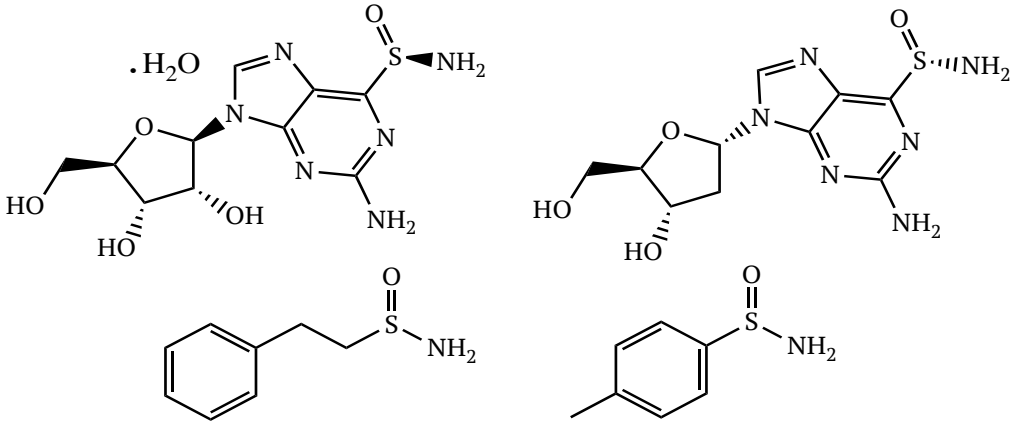


Figure 12: The molecular strucures of JAZZIT[74] (52, top left), VICGET[75] (53, top right), XEPHOR[76] (54, bottom left) [zzx01])and ZIJGOQ[1] (31Q, bottom right).

Despite the small quantity of literature available, these primary sulfinamides structures are likely to provide some, albeit limited, insight into the expected bonding motifs for the sulfinamides in this study. Of the two large ribonucleoside derivatives, the 3D co-ordinates of JAZZIT[74] [(*R,S*)-2-Amino-9-β-*D*-ribofuranosylpurine-6-sulfinamide hydrate (**52**, Figure 12, top left)] are undetermined and so the structure cannot be visualised using Mercury structure visualisation software.[45] VICGET[75] [(*S*)-2-Amino-9-(2-deoxy-α-*D*-erythro-pentofuranosyl)purine-6-sulfinamide (**53**, Figure 12, top right)] contains some

alternative hydrogen bond donors and acceptors and so the anticipated interaction between the primary sulfinamide moieties is disrupted in favour of alternative motifs, as has been displayed for primary amides above (section 2.3.2). The S=O hydrogen bond acceptor participates in two hydrogen bonds, one with the amine of the purine moiety creating a C(7) chain (2.2 Å), and another with a hydroxyl hydrogen, creating a C(11) chain (2.8 Å, Figure 13 and Figure 14). The N-H of the sulfinamide donates discrete hydrogen bonds to two different hydroxyl oxygens creating a complex $R_3^3(24)$ trimer at a binary level with the N-H···O-H bond (2.1 and 2.2 Å, Figure 15).

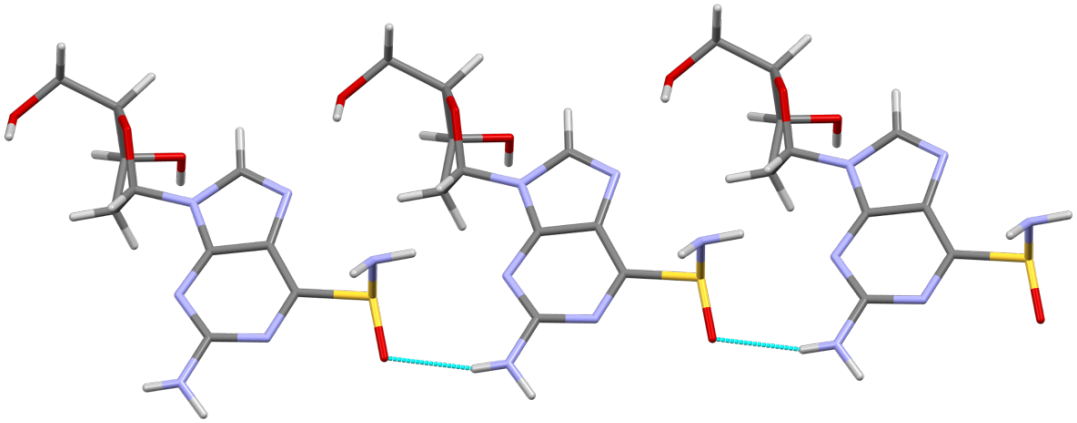


Figure 13: C(7) N-H···O=S chains in the crystal structure of VICGET[75] [53].

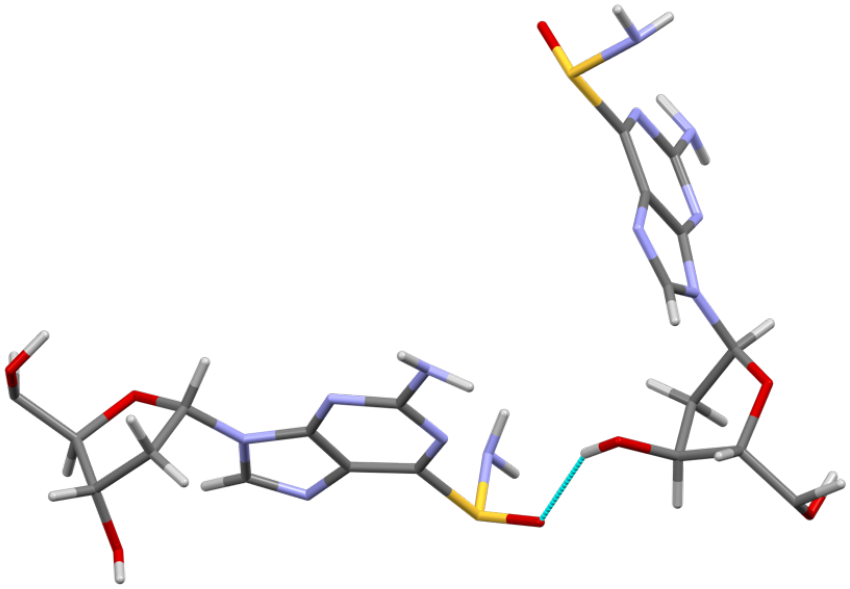


Figure 14: C(11) O=S···O-H chains in the crystal structure of VICGET[75] [53].

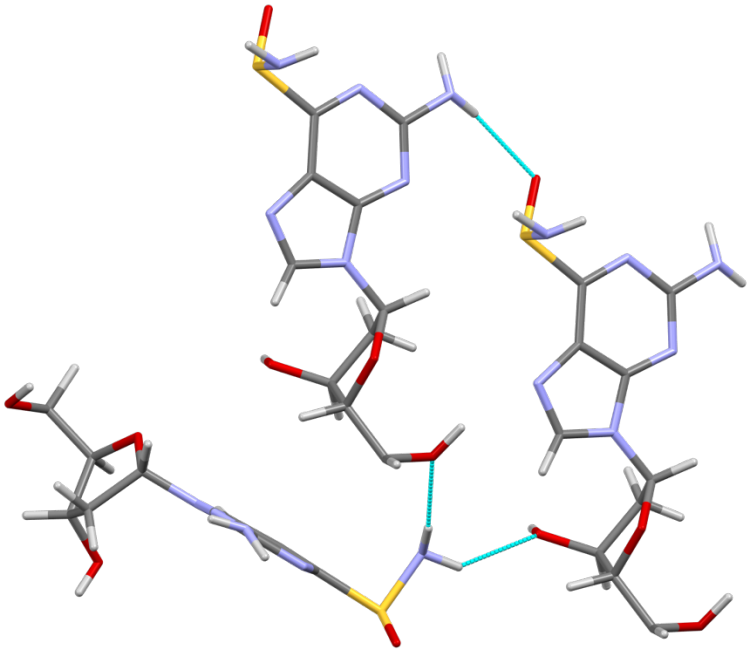


Figure 15: N-H···O-H and N-H···O=S interactions creating a trimer in the crystal structure of VICGET[75] [53].

These bonding features are interesting, in that they would not be traditionally predicted using Hunter's table[41] and Etter's rules.[77] The sulfinamide S=O (β = 8.3) is the most potent hydrogen bond acceptor within this structure, and yet does not accept hydrogen bonds from the most potent hydrogen bond donor, the sulfinamide N-H (α = 3.2). The sulfinamide S=O accepts two hydrogen bonds, one from a hydroxyl O-H (α = 2.7), and the other from the N-H on the pyrimidine (α = 1.5). The sulfinamide N-H donates its hydrogen bonds to two hydroxyl oxygens (β = 5.8). The multiple functionality of the molecule as a whole seems to strongly influence the hydrogen bonding abilities of the individual functional groups and a more close-packed arrangement supported by these weaker hydrogen bonds proves more favourable than the 'best donor to best acceptor' relationship. Nevertheless, even in a complex molecule such as VICGET[75] [**53**] with many competing donors and acceptors, the sulfinamide is proving to be strongly structure defining.

The two small molecule structures XEPHOR[76] [2-phenylethane sulfinamide (**54**, Figure 12, bottom left)] and ZIJGOQ[1] [4-methylbenzene sulfinamide (**31Q**, Figure 12, bottom right)] provide more valuable insight into the intrinsic interactions between sulfinamide moieties. Although the structural differences between these two molecules are modest, the solid state interactions display some distinct

differences. In the case of ZIJGOQ[1] [**31Q**], the repetitive structural motif is two infinite N-H···O=S C(4) chains (Figure 16), analogous to that which can be seen in many amide structures.[1] Described by Lawrence *et al.*,[1] the individual enantiomers of (±)-4-methylbenzene sulfinamide [**31Q**] interact with each other to form 'alternating layers in the *ac*-plane' (Figure 17).

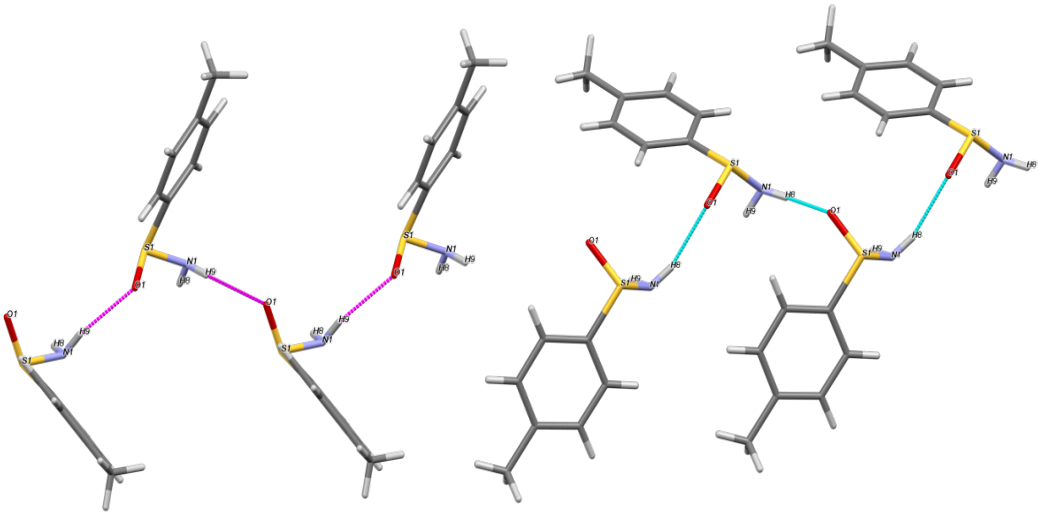


Figure 16 Two sets (left and right) of C(4) chains in 31Q [C(4) chains coloured in magenta and cyan, ZIJGOQ].[1]

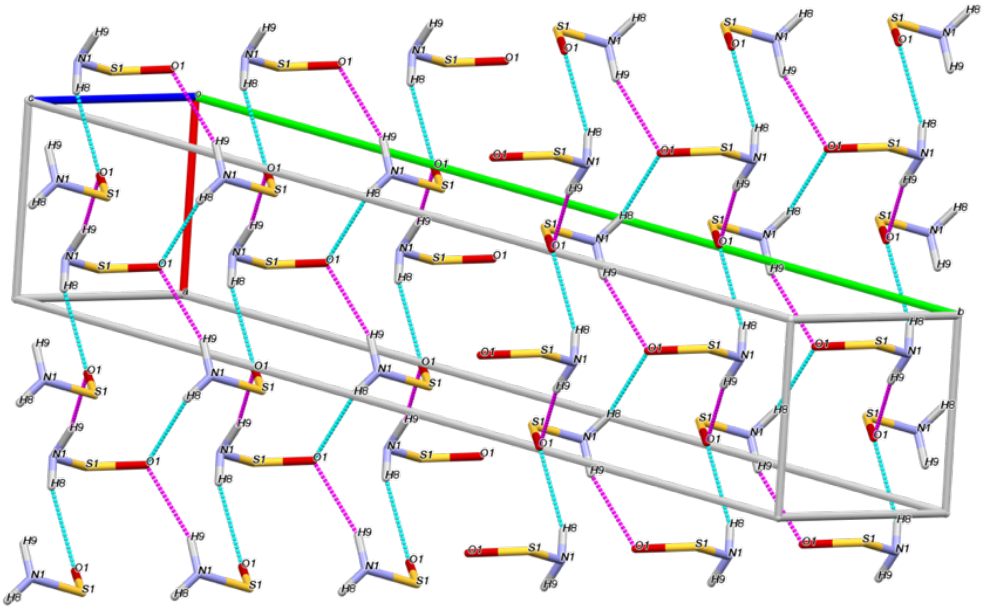


Figure 17: Alternating layers formed by *C*(4) chains in 31Q (aromatic rings removed) [C(4) chains coloured in magenta and cyan, ZIJGOQ].[1]

In contrast, the structure of XEPHOR[76] [**54**, 2-phenylethanesulfinamide] contains the C(4) chain motif at a unitary level, but at a binary level the combination of the two C(4) chains creates an $R_2^2(8)$ dimer-type interaction, comparable to the dimers observed in amides and carboxylic acids (Figure 18). Growth extends in a more linear fashion in comparison to **31Q**, creating an extended chain along the *b*-axis (Figure 19). The enantiopure form of **31Q** was also characterized and published as part of this study;[1] interestingly, the same pattern of alternating C(4) chains is observed in the enantiopure form of this material, with the orientation of the aromatic rings adjusted to support the overall motif.

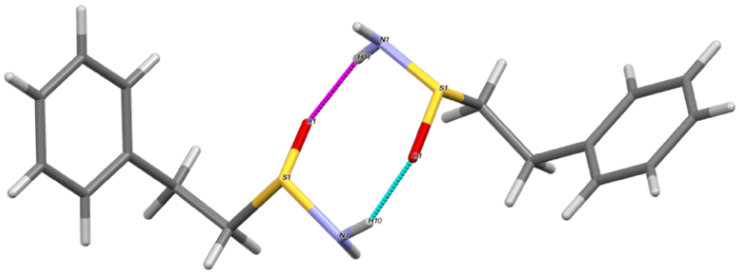


Figure 18: The $R_2^2(8)$ dimer-type motif in 2-phenylethane sulfinamide [54, XEPHOR].[76]

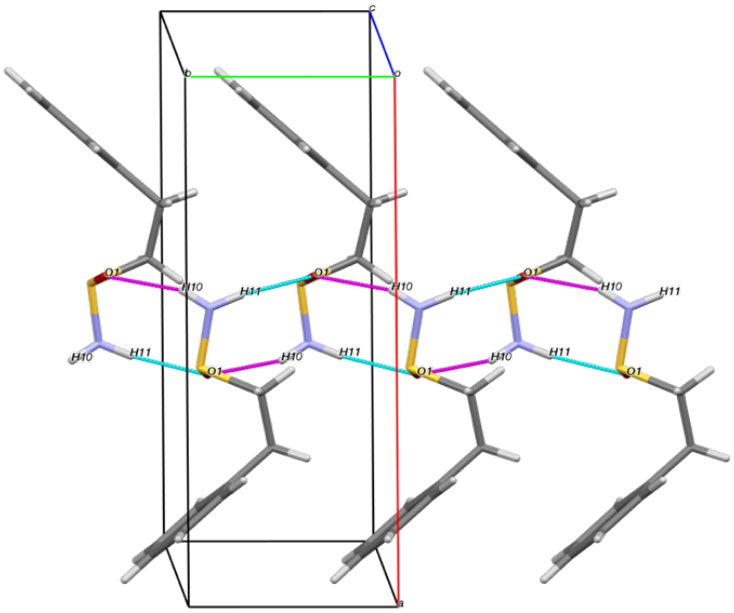


Figure 19: Extended *C*(4) chain along *b*-axis in 54 (XEPHOR).[76]

Overall, the analysis of these known structures supports the proposal that N-H···O=S bonding interactions in sulfinamides are likely to be structure defining with features reminiscent of those of the very well investigated amides i.e. the

C(4) chain and dimer-type interactions; although the exact structures are more complex, in part, due to the non-planarity at the sulfur and nitrogen atoms in the sulfinamide.

2.3.4 Conclusions

With only four examples of primary sulfinamides in the literature it is difficult to draw definite conclusions about the interactions to be expected in the solid state structures of primary aryl-sulfinamides [**31A, C-Q, S-T**]. Nevertheless, the following ideas can be utilised going forward:

(a) The sulfinamide moiety warrants study from a solid state perspective.
(b) Sulfinamides can be expected to crystallize exhibiting strong N-H···O=S hydrogen bonding interactions in the solid state.
(c) There is scope for variation within the N-H···O=S hydrogen bonds to create different supramolecular architectures.
(d) If alternative hydrogen bond acceptors are presented within the molecular structure, the N-H···O=S interaction can be disrupted to accommodate alternative hydrogen bonding networks including these acceptors.

In contrast to the planar amides, the intrinsic chirality in the sulfinamide moiety enables the study of chirality in the solid state. The specific focus of this work is to use the crystal structures of **31A, C-Q, S-T** to understand the nature of the N-H···O=S interaction between the sulfinamide groups, particularly in relation to the persistence of specific motifs in the presence of alternative hydrogen bond donors and acceptors; and secondly, to observe how the individual enantiomers of the racemic sulfinamide behave relative to each other in supporting the supramolecular architectures.

2.4 Crystal Structures of Sulfinamides [31A-T]

Crystallization of the 18 sulfinamide materials was required in order to prepare good quality crystals for single crystal analysis. The primary method of single crystal preparation was slow evaporation at room temperature from solutions of the sulfinamide in volatile organic solvent. Slow evaporations were performed from methanol, ethanol, isopropanol (IPA), acetonitrile, chloroform ($CHCl_3$) and CH_2Cl_2, where solubility would allow dissolution of the sulfinamide.

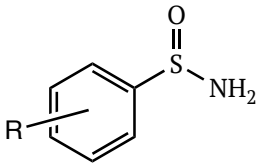
The crystallization process proved a major challenge when preparing these materials, with only three materials yielding crystals suitable for SCXRD at the first slow evaporation attempt [**31N** (4-MeO), **31S** (3-NO_2) and **31T** (4-NO_2)]. Ultimately, 11 of 18 sulfinamides yielded crystals that were successfully characterised using single crystal diffraction, and their bonding motifs in the solid state analysed. Overall, ethanol proved to be the most successful solvent for crystallization of these materials (9 of 11 crystal structures), but multiple attempts at crystallization were required in order to yield crystals suitable for SCXRD in 6 of these cases. It is reasonable to conclude that the experience gained in handling poor quality crystals as the project progressed contributed more to the determination of successful structures than the repetition of the crystallizations.

Unexpectedly, hydrolysis occurred commonly in the crystallization solvent, resulting in determination of an unanticipated crystal structure in 7 of the 19 materials (these results will be discussed in section 2.5). To counteract this issue, crystallization experiments were attempted in dry ethanol, under a nitrogen atmosphere in a dessicator over P_4O_{10} in order to minimise the presence of water. However, this method proved successful in yielding single crystals of the parent sulfinamide product for **31C** (2-Br) only. A mixture of IPA and $CHCl_3$ was used to yield single crystals of **31E** (4-Br), after attempts to crystallize from all of the above solvents had failed. **31M** (3-MeO) proved unstable with respect to hydrolysis in all solvents, but the crystal structure was finally determined from the neat reaction product without recrystallization. Following this success, attempts were made to choose suitable single crystals from the neat synthesised

powder for the remaining 7 compounds in the series, but no other crystal structures could be determined this way.

4-Methylbenzene sulfinamide (**31Q**, ZIJGOQ) was included in the series to investigate whether the bonding motifs displayed within a racemically synthesised material would display differences to that crystallized from a mixture of the individual enantiomers, as we had previously described.[1] It has been shown that materials can perform differently depending when crystallized from a mixture of the individual enantiomers or the racemically synthesised material.[78] However, the unit cell determination revealed the crystal structure to be the same as that found in the literature,[1] and so resources were not used to re-determine this structure. The structures obtained crystallized into a range of space groups (Table 8). The space group distribution is interesting because it highlights the variety of space group symmetries available to these structurally related molecules.

Table 8: Space group distribution in primary aryl sulfinamides



R =	Space Group	Crystallization Solvent	Crystal Habit	Crystal Colour
2-Br [31C]	*Pbca*	Dry Ethanol[a]	Plate	Colourless
3-Br [31D]	*Pbca*	Ethanol	Plate	Colourless
4-Br [31E]	*Pc*	IPA / $CHCl_3$	Thin plate	Colourless
3-Cl [31G]	*Pbca*	Ethanol	Brick	Colourless
3-F [31J]	*Pbca*	Ethanol	Plate	Yellow
4-F [31K]	*Pbca*	Ethanol	Plate	Yellow
2-MeO [31L]	*P6cc*	Ethanol[b]	Plate	Colourless
3-MeO [31M]	*R3c*	None	Needle	Orange
4-MeO [31N]	*Pbca*	Ethanol	Plate	Colourless
4-Me [31Q][1]	-	Ethanol	Plate	Colourless
3-NO_2 [31S]	*Pbca*	Ethanol	Needle	Orange
4-NO_2 [31T]	$Pna2_1$	Ethanol	Needle	Orange

a - Slow evaporation in dry ethanol over P_4O_{10}; b - Quick evaporation of on a warm clock glass;

In agreement with the literature information, the structural architecture in primary aryl sulfinamides is determined by combinations of N-H···O=S hydrogen bonds forming C(4) chains. C(4) chains, as mentioned earlier (Sect. 2.3.2) are a common target in crystal engineering, conferring directionality to the crystal growth.[51]

Sulfinamides **31C** (2-Br), **31D** (3-Br), **31G** (3-Cl), **31J** (3-F), **31K** (4-F), **31N** (4-MeO) and **31S** (3-NO_2) all crystallize in the *Pbca* space group. **31G**, **31J**, **31K**, **31N** and **31S** exhibit very similar supramolecular architectures with the same spatial pattern of C(4) chains, comprised of alternating *R*- and *S*- enantiomers. These chains combine at the binary level to create a spiral down the *a*-axis, cross-linked with C(4) chains along the *b*-axis (Figure 20 and Figure 21). The supramolecular system creates hydrophilic and hydrophobic areas within the structure (Figure 22).

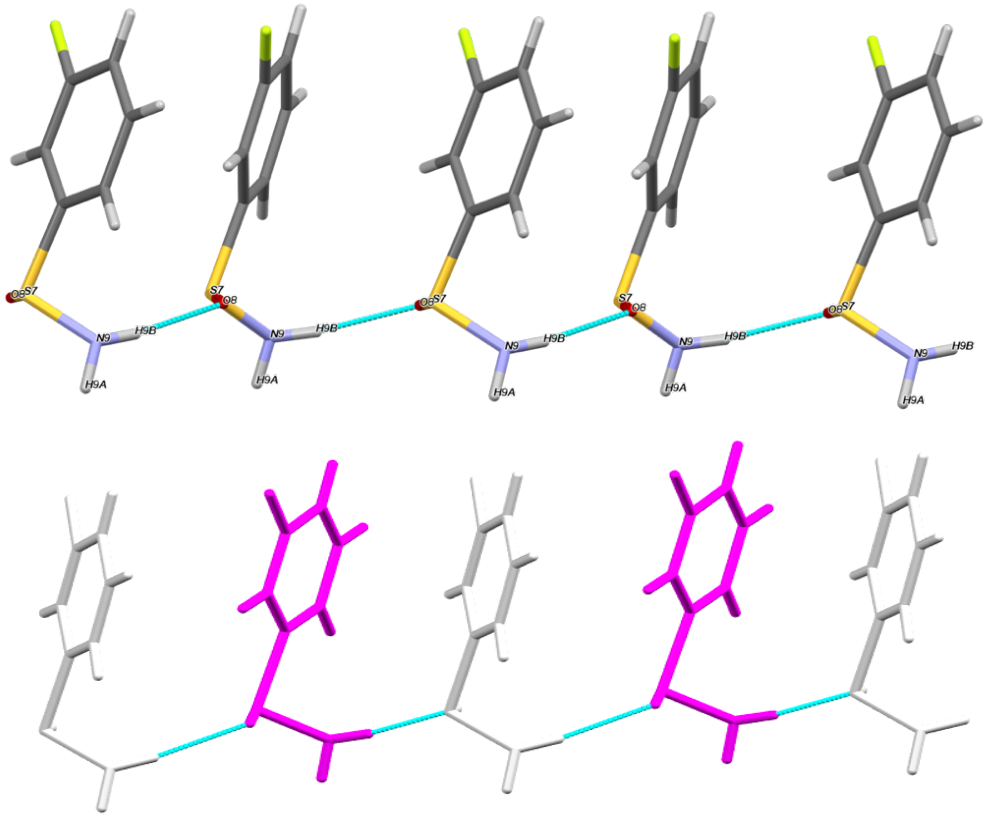


Figure 20: N-H_{9B}···O=S interaction, forming C(4) chains along the *a*-axis in 31J (3-F) [top], enantiomers coloured magenta and white as related by glide plane [bottom], hydrogen bonds in cyan.

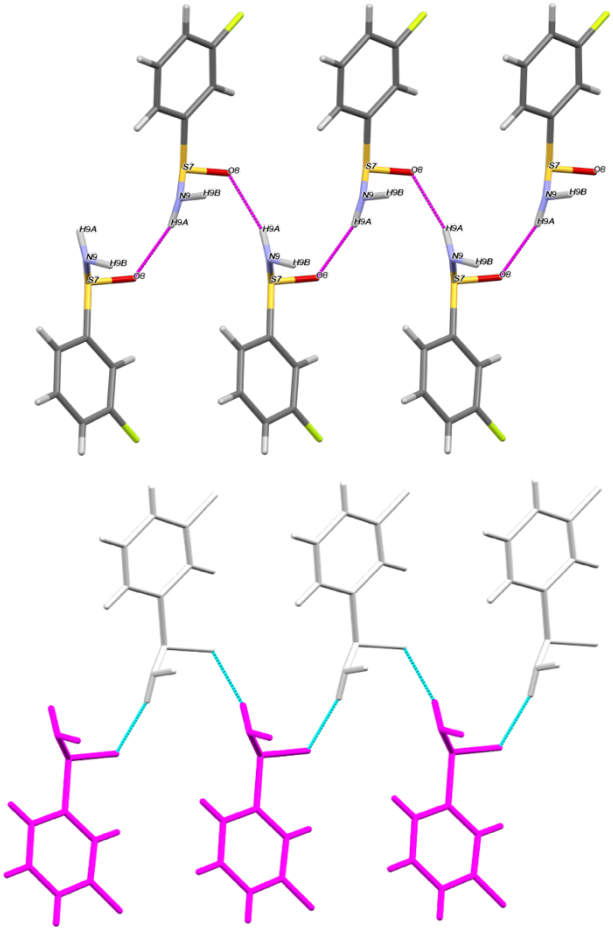


Figure 21: N-H_{9A}···O=S interaction, forming C(4) chains along the *b*-axis in 31J (3-F), enantiomers coloured magenta and white as related by glide plane [bottom].

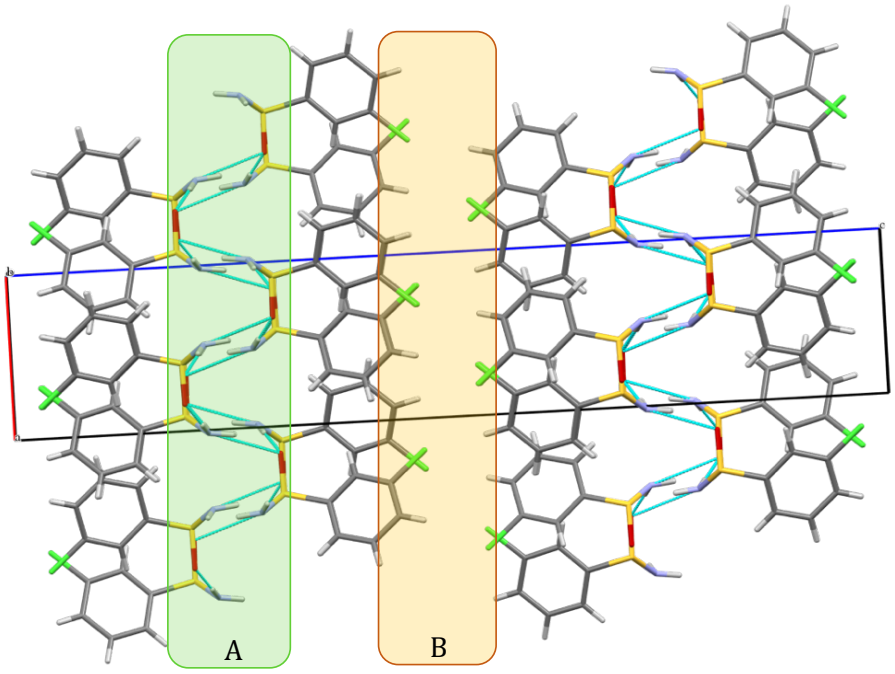


Figure 22: Formation of hydrophilic (A, green) and hydrophobic (B, orange) areas within the crystal structure of 31G (3-Cl).

The 4-methoxy derivative [**31N**] presents additional weak hydrogen bonds within the hydrophobic areas of the structure. There is a C-H···O-C interaction between methyl groups on adjacent molecules that create $R_2^2(6)$ dimers that further stabilise the structure (Figure 23 and Figure 24). This type of C-H···O-C interaction has received considerable debate in the literature,[79] particularly in relation to its validity as a true, structure defining bond as opposed to a structural artefact arising as a consequence of the dominant hydrogen bonding motifs present in the structure. If a conclusion to this debate were to be drawn from this structure alone, it would be fair to say that the occurrence of the C-H···O-C interaction in this structure is likely a structural artefact; based upon the similarity of the predominant structural motifs observed here with the other sulfinamide structures (above).

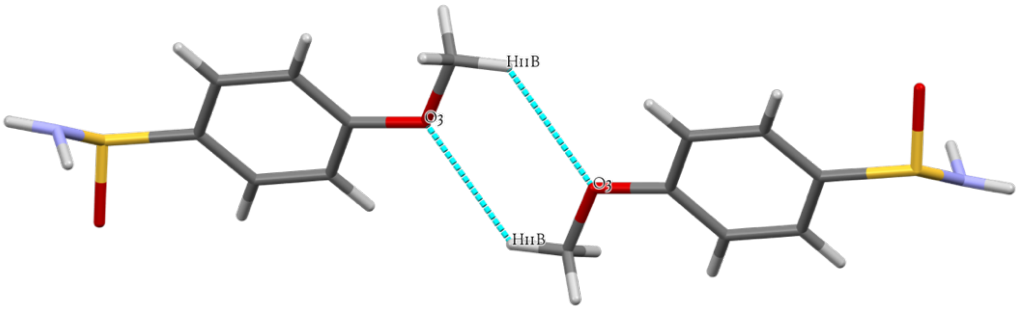


Figure 23: The $R_2^2(6)$ dimer formed in 31N (4-MeO).

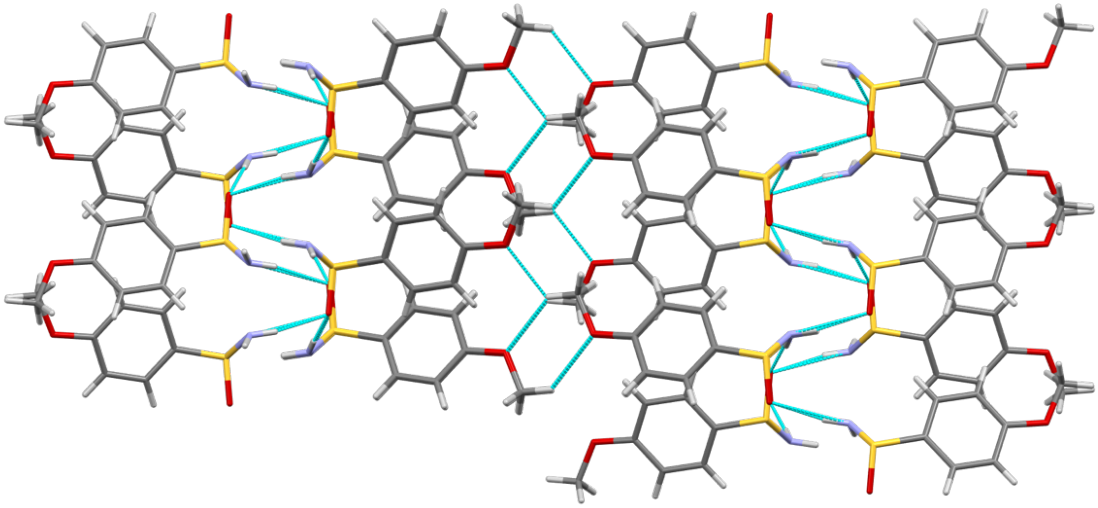


Figure 24: The hydrophyllic and hydrophobic region formed in 31N (4-MeO).

A very similar pattern of hydrogen bonds can be observed in the structure of 4-methylbenzene sulfinamide (**31Q**, ZIJCOQ, $P2_1/n$) as determined previously within our research group.[1] C(4) chains create a similar spiral pattern down the *c*-axis, and extend along the *b*-axis (Figure 25), resulting in alternating layers, as discussed above (Figure 16 and Figure 17). The crystal structure of enantiopure (*S*)-4-methylbenzene sulfinamide (**(*S*)-31Q**, ZIJGUW) was also published as part of this work, displaying a similar pattern of hydrogen bonds, facilitated by the orientation of the aromatic rings.

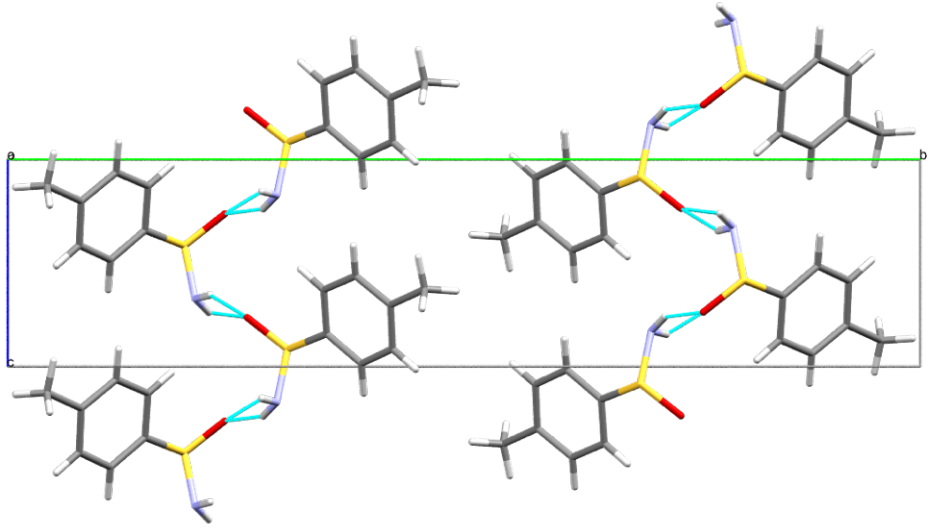


Figure 25: C(4) chains spiralling down the *c*-axis in ZIJCOQ (31Q, 4-Me).[1]

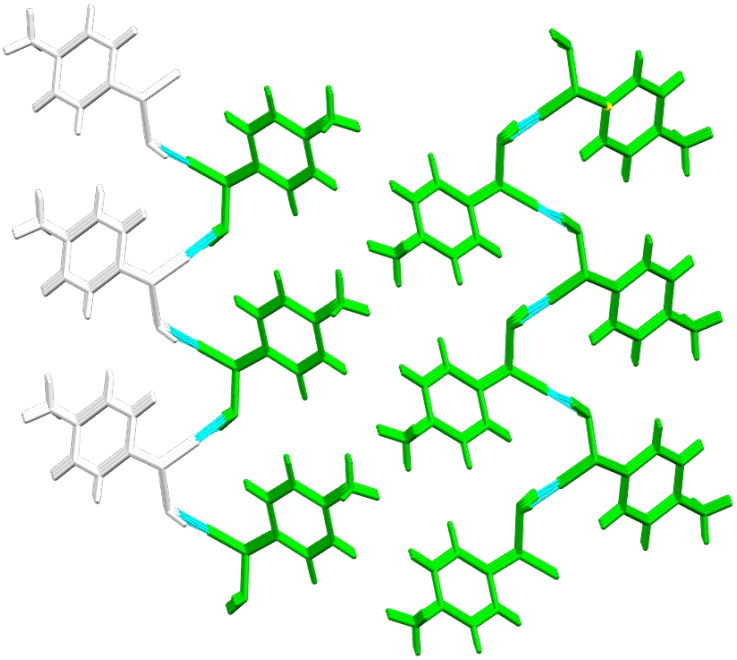


Figure 26: Hydrogen bonding pattern in (*S*)-4-methylbenzene sulfinamide ((*S*)-31Q ZIJGUW); white to green as related by 2_1 screw axis.[1]

In contrast to the amides discussed earlier, where the presence of the nitro group led to disruption of the C(4) chain, the corresponding nitro-sulfinamides **31S** and **31T** retain their C(4) chain motifs, indicating that the S=O···H-N interaction is very robust. The 3-nitro substituted material (**31S**), displays the C(4) chain motifs shown above (Figure 27), with additional interactions between the nitro groups in the structure also observed.

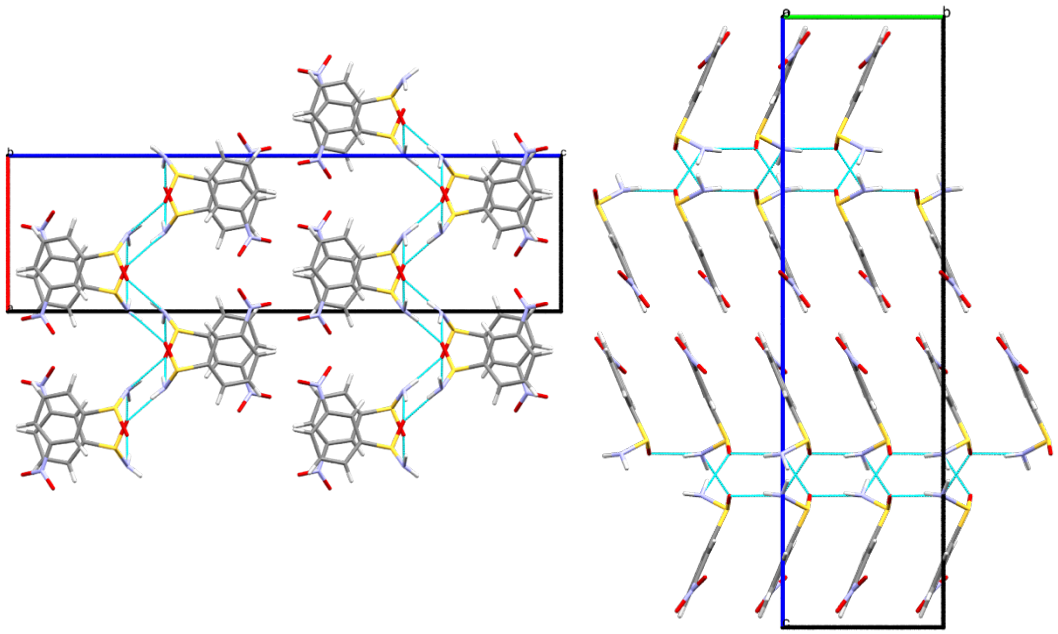


Figure 27: C(4) chain interactions in 3-nitrobenzene sulfinamide (31S).

31S displays π-π stacking of N=O groups at a length of approx. 3.1 Å, and stacking of the aromatic rings at approx. 3.9 Å, along with a C-H···O=N interaction at a length of approx. 2.5 Å. This facilitates the formation of layering of the molecules overall (Figure 28), as the additional interactions hold the phenyl ring in a planar orientation to the sulfinamide S=O, with the nitro group twisted only slightly at 15° (Figure 28).

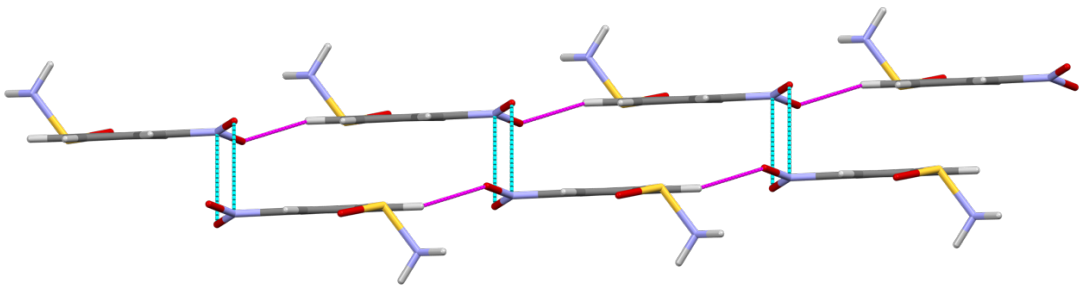


Figure 28: π-π stacking (cyan) and C-H···O=N interactions (magenta) in 31S (3-NO_2).

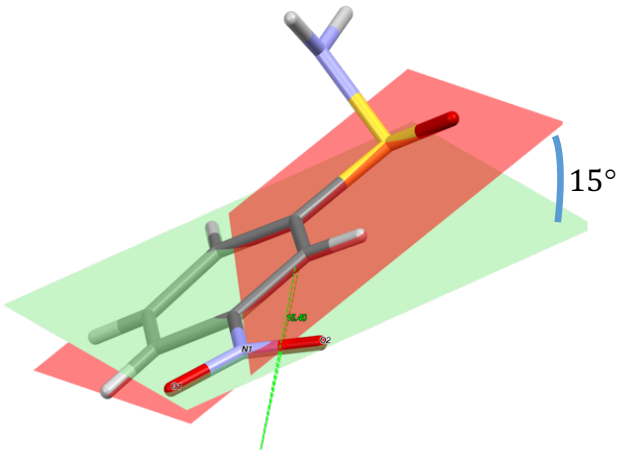


Figure 29: Planarity of the phenyl ring and slight twist of NO_2 group in 31S (3-NO_2).

In contrast, in the case of 4-nitrobenzene sulfinamide [**31T**] (*Pna*2_1), the C(4) chain is also the dominant motif, while the overall architecture is quite different. The chains of alternating enantiomers adopt a linear configuration extending crystal growth along the *c*-axis and spiralling down the *b*-axis (Figure 30 and Figure 31).

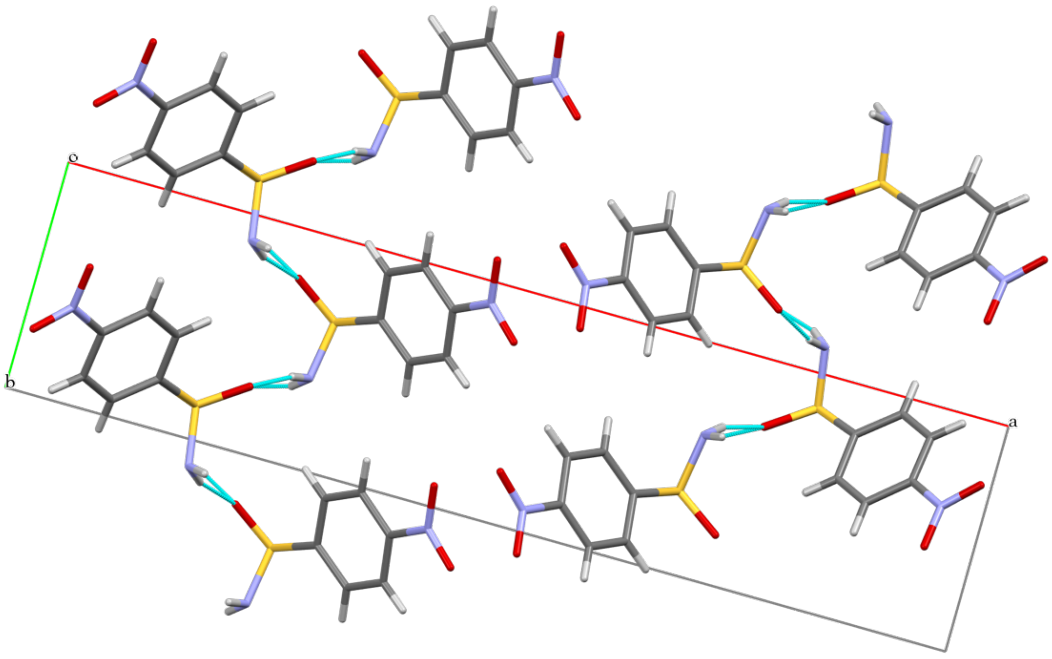


Figure 30: *C*(4) chains extending along the *b*-axis in 31T (4-NO_2).

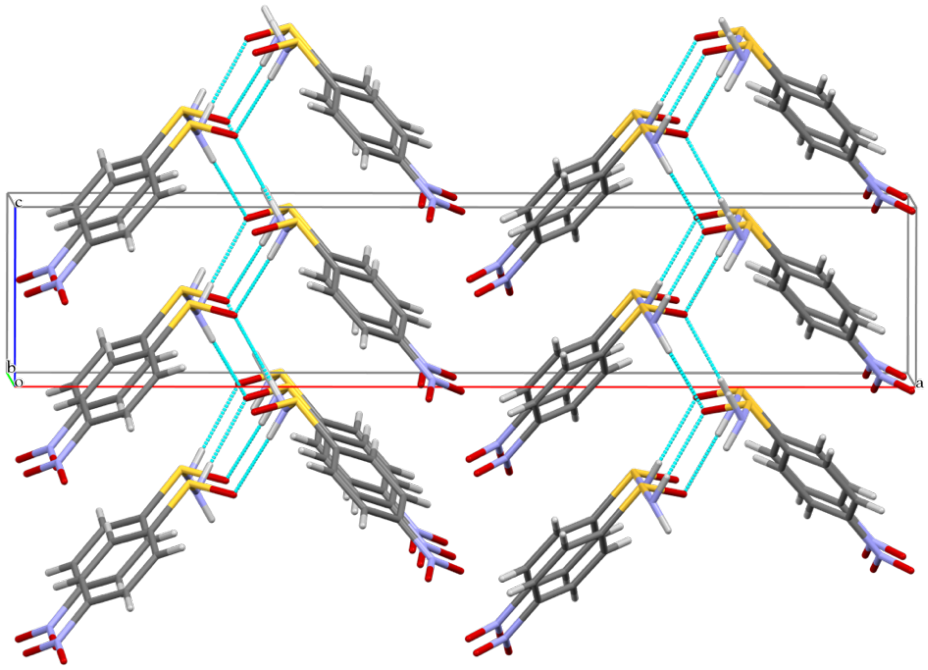


Figure 31: View along the *c*-axis in 31T (4-NO_2).

Chains between the nitro groups occur in place of the stacking interactions seen in the previous 3-nitro-substituted sulfinamide [**31S**]. The nitro group is also more planar, in relation to the phenyl ring, with an angle of just 4.4° between the two in this case. C-H···O=N interactions are observed also, creating almost perfectly perpendicular layers of the molecules at an angle 87.4° (Figure 32). This almost perpendicular relationship is a common feature of nitro-nitro interactions in the solid state.[80] The relative position of the nitro group precludes the 3-substituted analogue from adopting such an angular geometry, but conversely the 4-nitro material could not access the layered motifs displayed by the 3-nitro analogue due to the sulfinamide N-H···O=S interactions taking precedence over the weaker nitro-nitro interactions. The combination of these interactions creates a distinct zig-zag motif at a supramolecular level, with the aromatic rings parallel to each other in the respective layers (Figure 33). In contrast with amides, where the presence of the nitro group disrupts the primary hydrogen bonding pattern, the resilience of the sulfinamide N-H···O=S interactions is striking.

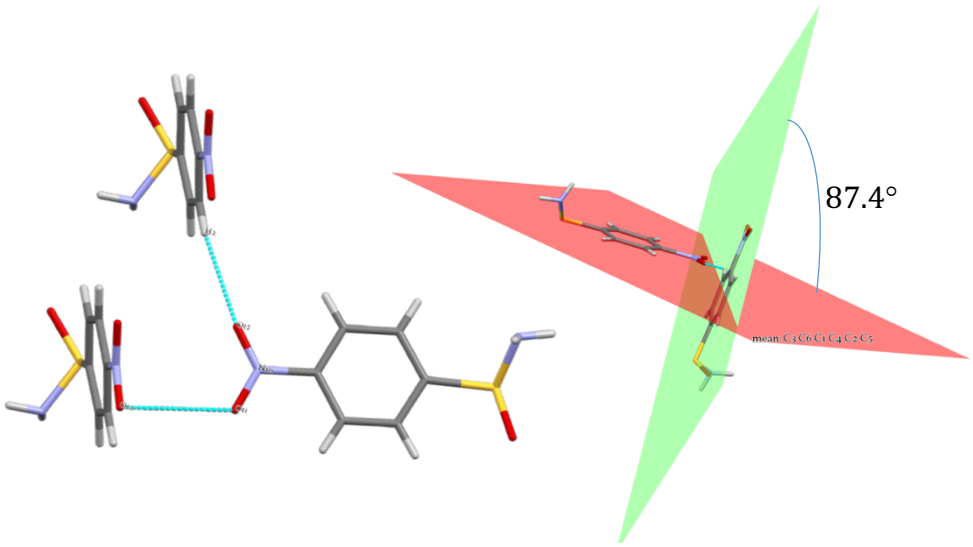


Figure 32: N=O···O=N and C-H···O=N interactions in 31T (4-NO_2) [left] leading to perpendicular geometry.

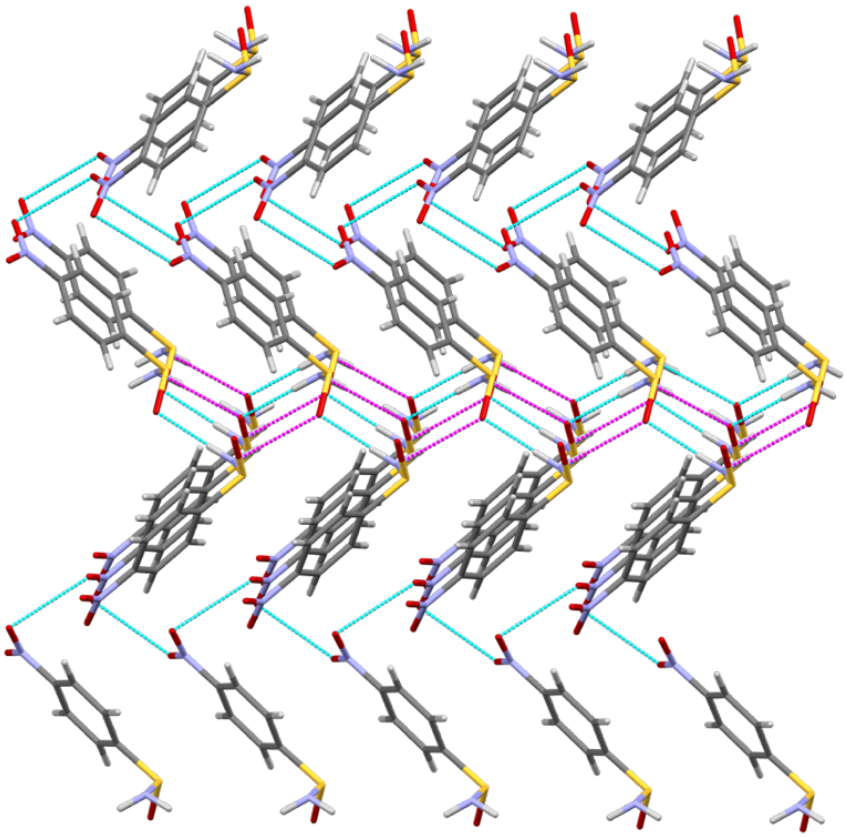


Figure 33: Zig-zag layers of molecules in 31T (4-NO_2).

The 2-bromo derivative [**31C**] also exists in the *P*bca space group, but displays distinct differences when compared to those discussed above. There are two separate C(4) chain interactions, the first along the *b*-axis, which is comprised of alternating enantiomers (Figure 34); and the second, along the *a*-axis, comprised of single enantiomers (Figure 35).

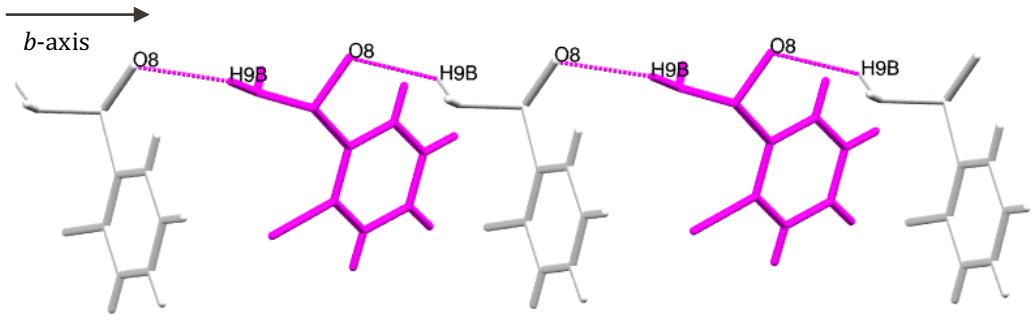


Figure 34: C(4) chains of alternating enantiomers along the *b*-axis in 31C (2-Br) [magenta] (molecules related by glide plane white to magenta).

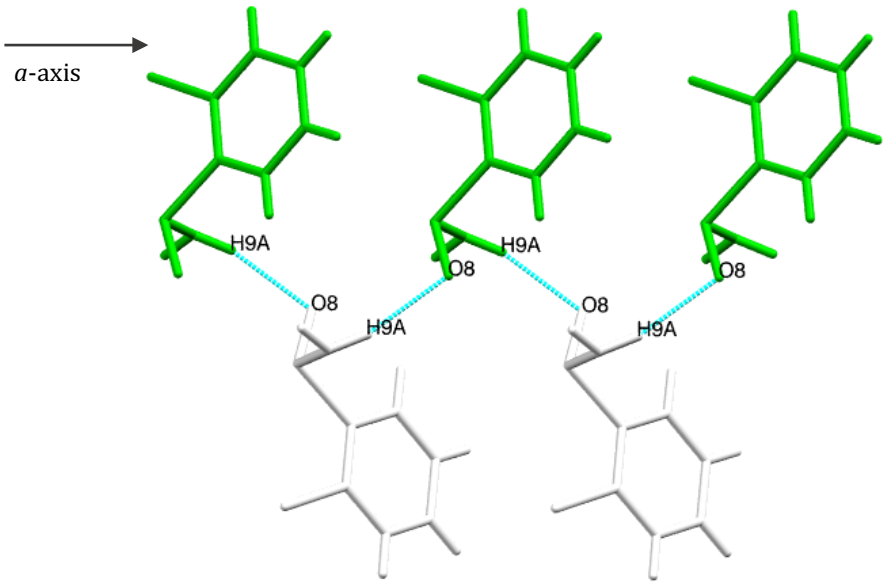


Figure 35: Enantiopure C(4) chains along the *a*-axis in 31C (2-Br) [cyan] (molecules related by 2_1 screw axis (white to green).

These two interactions crosslink at a binary level to form an $R_4^2(8)$ tetramer (Figure 36). The distinct linearity of the interactions creates thin, plate crystals that are very challenging for structure determination from SCXRD.

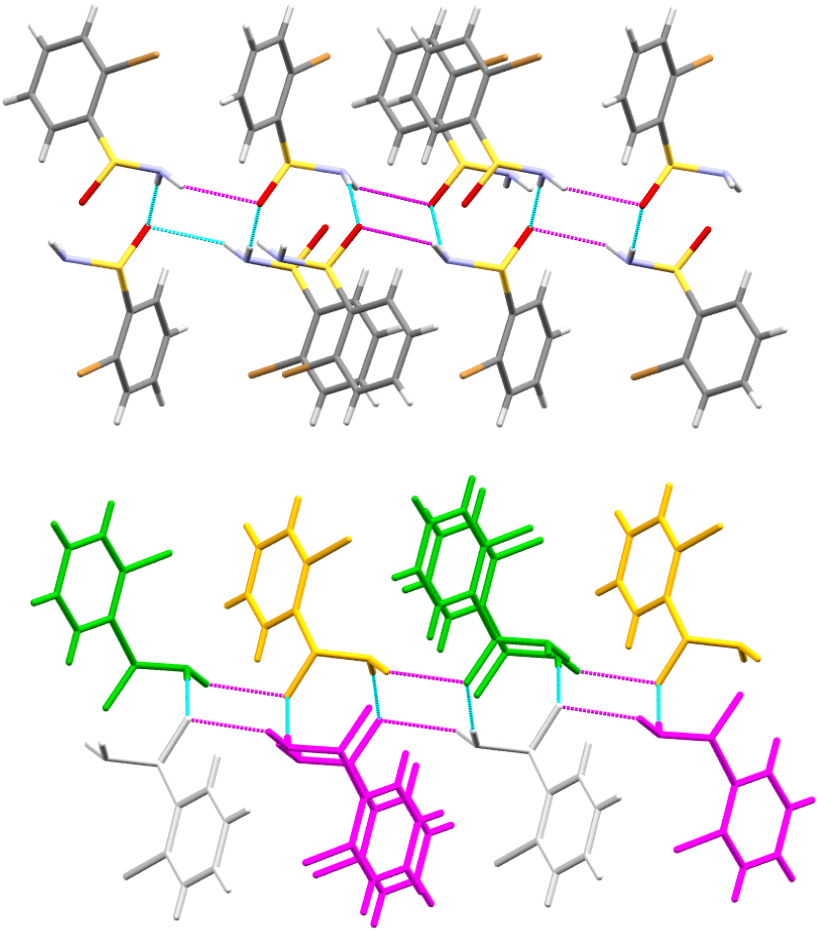


Figure 36: $R_4^2(8)$ tetramer formed by crosslinking of C(4) chains in 31C (2-Br) [top] and symmetry operations [bottom] (hydrogen bonds in cyan and magenta).

A similar, but greater challenge to that encountered in crystallizations of **31C** was presented by its 4-bromo counterpart [**31E**]. Single solvent solution crystallization of this material was attempted from every solvent to hand that dissolved the material, however, in all instances the material crystallized in extremely thin, film-like crystals that provided little to no diffraction in SCXRD. Ultimately, successful structure determination (R factor approx. 9%) was performed on this material as crystallized from a mixture of $CHCl_3$ and IPA after several attempts. The crystal structure was determined in the space group P*c*.

The hydrogen bonding motifs observed in this material are quite different to that observed in any of the other materials, because two crystallographically unique molecules exist in the structure, each creating C(4) chains of alternating enantiomers along the *c*-axis (Figure 37).

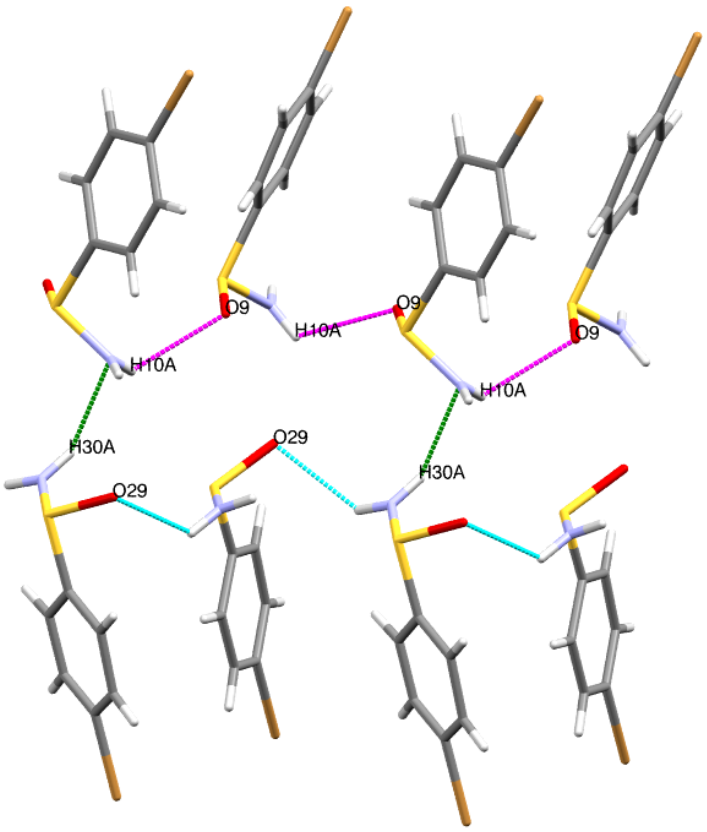


Figure 37: Two sets of C(4) chains along the *c*-axis in 31E (4-Br), C(4) chains in cyan and magenta, H_{30A}···N-H interaction in green.

One of the most interesting features of the bonding in this structure is that only one of the two N-H interacts in hydrogen bonding chains for each unique molecule. H_{30A} points towards the nitrogen lone pair of the nearest sulfinamide moiety at a distance of approximately 2.2 Å (Figure 37), but does not appear to participate in any strong hydrogen bonding interactions in the way that has been observed for all other sulfinamides thus far. H_{10B} does not appear to point toward any centre of electron density. These observations could be due to the lack of appropriate acceptors within the acceptable distance around the hydrogen atoms. The closest hydrogen bond acceptors (S=O) in the structure are approximately 3.2 Å and 3.8 Å away from H_{30A}, and 2.8 Å and 4.5 Å away from H_{10B}. This lack of hydrogen bonding interactions is a potential explanation for the poor crystallinity observed when working with this material. PXRD analysis was performed on the sulfinamide product, and it was observed that the material was largely amorphous, with only a small amount of crystalline content.

The interactions observed in this material can be used to rationalise the morphology of the crystals. There is no expansion observed along the *a*-axis in

this crystal structure, leading to crystal growth along two axes with a lack of hydrogen bonding interactions that would give depth to these crystalline sheets. A similar crystal morphology was observed in the case of the 4-chloro analogue of this material [**31H**]. Unfortunately, attempts to determine the crystal structure of this compound were unsuccessful as the thin sheets diffracted very poorly. It is likely that the crystal structure of the 4-chloro analogue [**31H**] displays a similar structural architecture to that observed in **31E** (4-Br), with the bromo-substitution providing enough localised electron density to allow successful structure determination.

It is intriguing that replacing the 4-fluoro with 4-bromo or 4-chloro results in a dramatic alteration in the solid state properties, presumably due to the steric effects of the electron rich 4-halo substituents. Analysis of the interfluoro interaction distance in **31K** shows that the shortest intermolecular F-F distance is in the order of 3.1 Å (Figure 38). Clearly, replacement of the fluorine substituent with a larger chlorine or bromine atom would not be feasible while retaining the same solid state architecture.

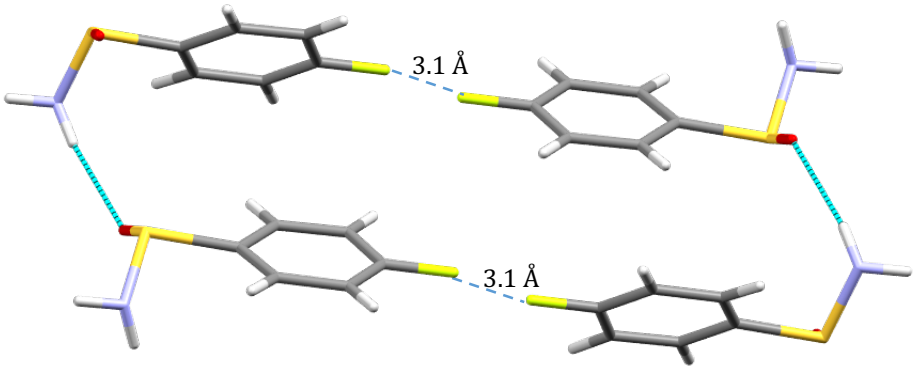


Figure 38: Intermolecular F-F distance in 4-fluorobenzene sulfinamide (31K).

The final two structures determined in this series were the 2- and 3-methoxy substituted sulfinamides **31L** and **31M**. In the case of **31L** (*P*6cc), in order to obtain single crystals and prevent hydrolysis, a small amount of the product was dissolved in ethanol, but instead of the traditional slow evaporation method, the solution was allowed to evaporate quickly on a clock glass which had been pre-warmed in the oven.

The primary interaction observed here is the usual C(4) chain of alternating *R*- and *S*- enantiomers along the *c*-axis, with only one of the two donor hydrogens [N-H_{11B}] involved in the formation of this chain (similar to that seen above for **31E**). The C(4) chains combine in groups of three at an architectural level to form large ring structures around the 3-fold rotation axis in the hydrophilic areas of the structure (Figure 39) and secondly, around the 6-fold rotation axis in the hydrophobic areas (Figure 40).

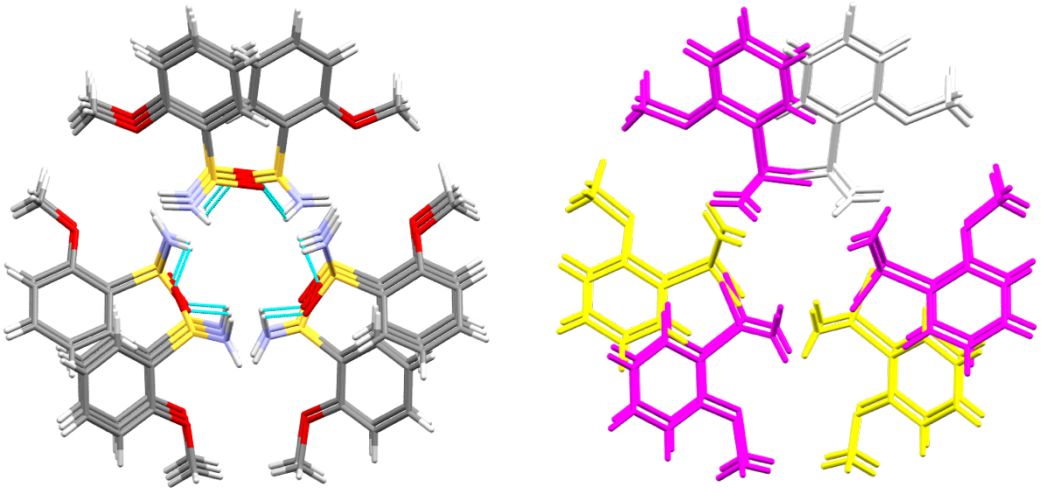


Figure 39: Ring structure formed by 3 C(4) chains in 31L (2-MeO) [view down *c*-axis], white to yellow *via* rotation, white to magenta *via* glide plane (therefore, white and yellow are one enantiomer, and magenta are the opposite enantiomers).

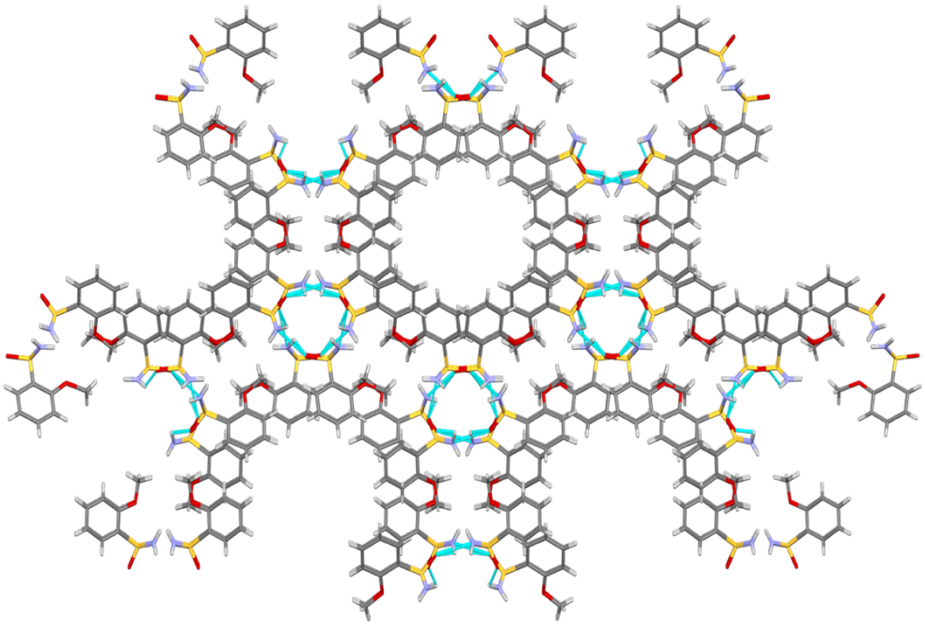


Figure 40: The overall crystalline growth pattern around the 3- and 6-fold rotation axes in 31L (2-MeO) [view down *c*-axis].

The crystal packing observed in this material creates hydrophilic channels of approximately 4.6 Å in diameter around the 3-fold rotation axis, within which

water can become trapped, while hydrophobic channels around the 6-fold axis are approximately 8.5 Å in diameter. In fact, a small quantity of non-stoichiometric water was observed within the hydrophilic channels in this structure that was removed using the SQEEZE function within the PLATON software.[81] The facile inclusion of water within the crystal structure of this material is considered to be a contributing factor to the relative ease with which the hydrolysis of this material can occur (Section 2.4).

The crystal structure of **31M** (3-MeO) was determined from neat crystals of the initial reaction product, therefore, no evidence of included water was observed in the structure. In a similar fashion to **31L**, **31M** has one-dimensional hydrophilic channels formed within the structure created by three interlinked C(4) chains with alternating enantiomers. Both sulfinamide hydrogen atoms participate in the formation of the linked C(4) chains in this structure, creating the ring around the 3-fold rotation axis and down the *c*-axis (Figure 41). Since this material crystallizes in the *R*3*c* space group, the hydrophobic channels in **31L** are not observed, although there are hydrophobic areas formed in the structure around the 3-fold rotation axes (Figure 42). In cases where **31M** was crystallized from single solvent slow evaporation, the hydrolysis product was observed in all cases, most likely due to the ease of introduction of water into these channels.

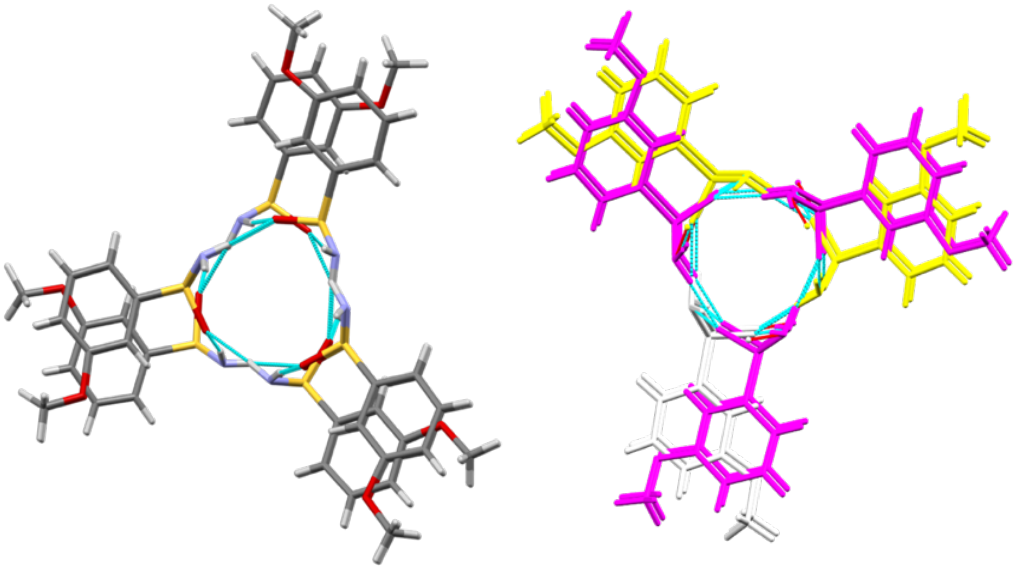


Figure 41: Ring structure formed around 3-fold rotation axis in 31M (3-MeO), white to yellow *via* rotation, white to magenta *via* glide plane (therefore, white and yellow are one enantiomer, and magenta are the opposite enantiomers).

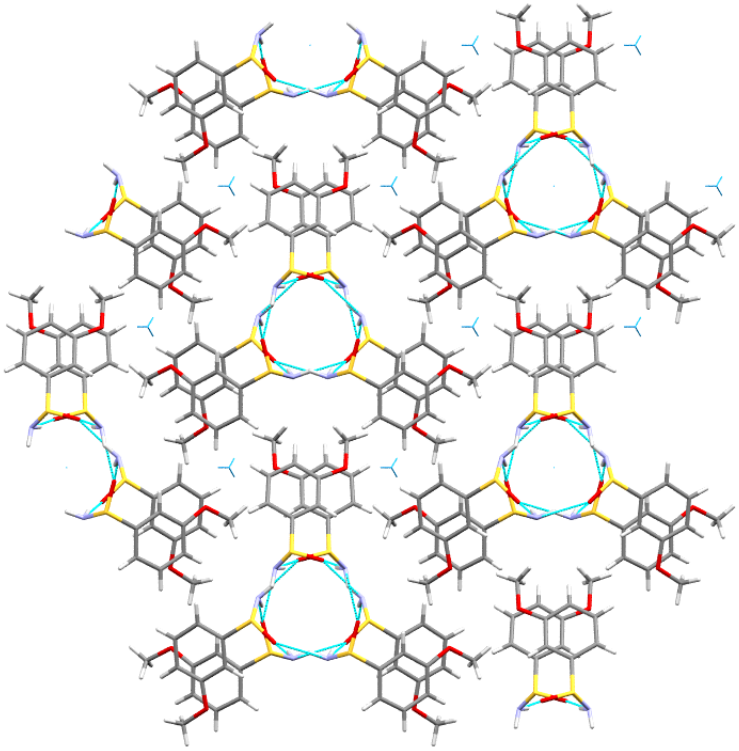


Figure 42: The overall crystalline growth pattern around the 3-fold rotation axes (blue) in 31M (3-MeO) [view down *c*-axis].

2.4.1 Conclusions

The 11 crystal structures determined in this project display the expected motifs of strong N-H···O=S interactions, creating directional C(4) chains that combine at a binary level to form a variety of interesting supramolecular architectures, including racemic and enantiopure motifs. Interestingly, the structures of the new racemic sulfinamides determined in this study are very similar to that previously reported within our research group for 4-methylbenzene sulfinamide (**31Q**, ZIJCOQ) [Figure 16 and Figure 17],[1] indicating that the N-H···O=S motifs are very robust and persist irrespective of the substituent on the aryl ring. In contrast to the motifs observed with the amide moiety, the N-H···O=S interactions observed in these materials seem to be stronger than the corresponding N-H···O=C hydrogen bonds; for example, the N-H···O=S interaction is retained in the presence of the nitro group. Furthermore, the crystal structure of (*S*)-4-methylbenzene sulfinamide [**31Q**, ZIJGUW],[1] while different due to its enantiopurity, retains a very similar pattern of the N-H···O=S hydrogen bonds, confirming the structure determining nature of this interaction.

Successful crystallization of these materials from single solvent solution evaporation techniques presented a significant challenge to this work, with many of the crystals either reacting to form hydrolysis products (discussed below), or creating poor quality crystals.

The ultimate objective of this project would be to explore cocrystallization of the sulfinamide moiety. Upon completion of this part of the project, however, it has become clear the hydrolytic sensitivity of the primary sulfinamides would complicate a cocrystallization investigation. This study has formed an excellent basis for future investigation into the cocrystallization behaviours of sulfinamides.

2.5 Sulfinamide Hydrolysis

2.5.1 Introduction to Sulfinamide Hydrolysis.

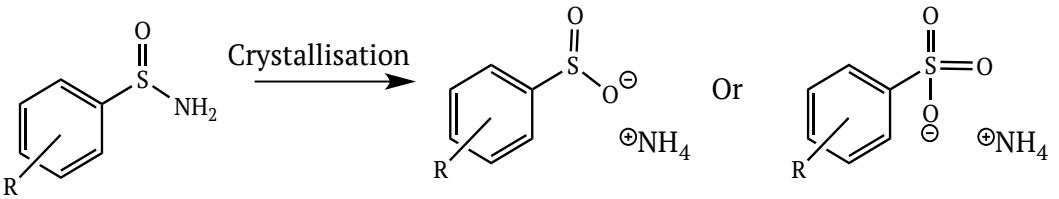
As mentioned previously, the hydrolysis of the sulfinamide materials in the crystallization solvent proved to be the major challenge in determining the crystal structures of the sulfinamide materials synthesised in this study. This unanticipated hydrolysis lead to a separate set of crystal structures and products; 7 of the 18 sulfinamide crystallizations produced either the ammonium sulfinate or sulfonate salt, rather than the parent sulfinamide material.

For sulfinamides **31D**, **31E**, **31F**, **31G**, **31J**, **31K**, **31N**, **31Q**, **31S** and **31T**, there was no evidence of hydrolysis observed in the crystallization experiments. As a result, it seems that some, but not all, sulfinamides are susceptible to hydrolysis in the organic crystallization solvent.

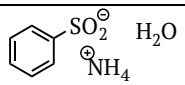
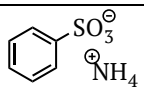
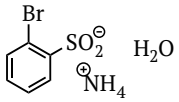
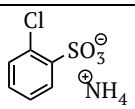
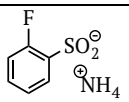
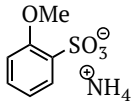
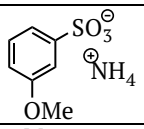
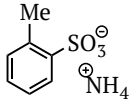
The experiments yielded crystal structures of two forms of the unsubstituted phenyl derivative (**55A** and **56A**), 5 *ortho*-subsitituted materials, and 1 *meta*-substituted material. Hydrolysis of the materials during attempted crystallization experiments proved reproducible, and when air was excluded from the crystallization of **31C** (2-Br), the pure sulfinamide was recovered. The crystal structure landscape of the ammonium sulfinate and sulfonate salts determined in this study is shown in Table 9.

The first crystal structure that was determined was **56F**, using crystals recovered from dissolution of the sulfinamide [**31F**] in acetonitrile for mass spectrometry studies, with other materials discovered periodically throughout the attempts to crystallize the sulfinamides.

Table 9: Crystal structures and solvents used for crystallization of sulfinate/sulfonate salts.



31A, C-Q, S-T **55A, 55C & 55I** **56A, 56F, 56L, 56M & 56O**

R =	Structure	Space Group	Solvent	Habit	Colour
H [55A]		*Pbca*	Dry Ethanol[a]	Block	Colourless
H [56A]		*Pbca*	Methanol	Plate	Orange
2-Br [55C]		$P2_1/c$	Acetonitrile	Plate	Colourless
2-Cl [56F]		$P2_1/c$	Acetonitrile	Block	Yellow
2-F [55I]		$P2_1$	Ethanol	Plate	White
2-MeO [56L]		*Pc*	Ethanol	Needle	Colourless
3-MeO [56M]		$P2_1/c$	Ethanol	Plate	Orange
2-Me [56O]		$P2_1/c$	Ethanol	Plate	Yellow

a - Crystallization performed in dry ethanol in a dessicator over P_4O_{10};

NMR characterisation of **56F** was performed in DMSO-d_6, as had been performed for the parent sulfinamides. However, it was observed that the signal for the ammonium counter-ion appeared in the 6-8 ppm region, overlapping with the signals for the aromatic protons. This meant that appropriate assignment of the C-H and N-H signals was very difficult. In order to complete the characterisation,

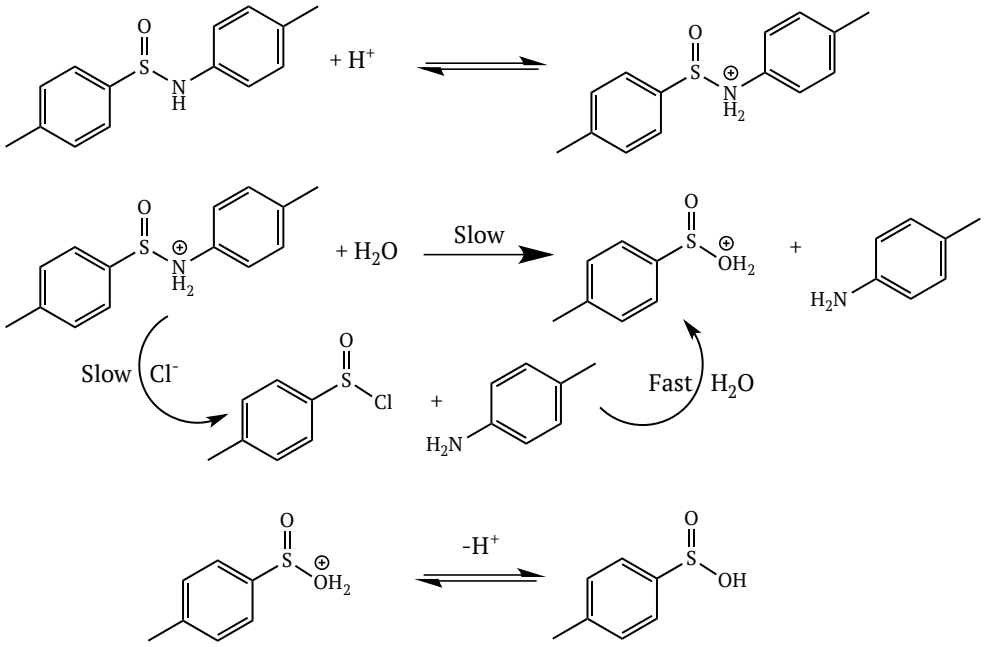
NMR studies of **56F** and subsequent sulfinate/sulfonate salts were performed in deuterated methanol; which lead to proton exchange with the ammonium and so the signal for the counter-ion was not observed, despite successful characterisation of the aromatic C-H signals. Deuterated methanol was used as the NMR solvent for characterisation of the remaining sulfinate/sulfonate materials. A large downfield shift was observed in the ^{13}C NMR spectrum for the carbon bearing the sulfur (8-9 ppm), indicative of the de-shielding of this carbon caused by the molecular change from sulfinamide to sulfinate/sulfonate salt.

Examination of the literature shows that the first reported study of sulfinamide hydrolysis was conducted by Biasotti and Anderson in 1971.[82] Hydrolysis of 9 *meta*- and *para*- substituted *N*-mesitylbenzene sulfinamides was conducted in basic aqueous ethanol. Results of this work determined that the reaction was first order in sulfinamide and first order in base (hydroxide).[82] Resonance effects were investigated in this instance, but no resonance effects were observed for the *para*-substituted arenesulfinamides (with the exception of *p*-methoxy, which gave a value just incrementally outside the experimental error). The work concluded that the lack of a resonance contribution for the *p*-nitro sulfinamide was a strong indication that there is no addition intermediate in the hydrolysis mechanism.

In 1977, Davis *et al*.[31] reported the formation of the corresponding sulfinic acid by refluxing 3-nitrobenzene sulfinamide [**31S**] in aqueous THF for 36 hours (yield 65-70%), and described the preferential formation of the hydrolysis product over the methanolysis product when the sulfinamide was refluxed in methanol for 7 days.

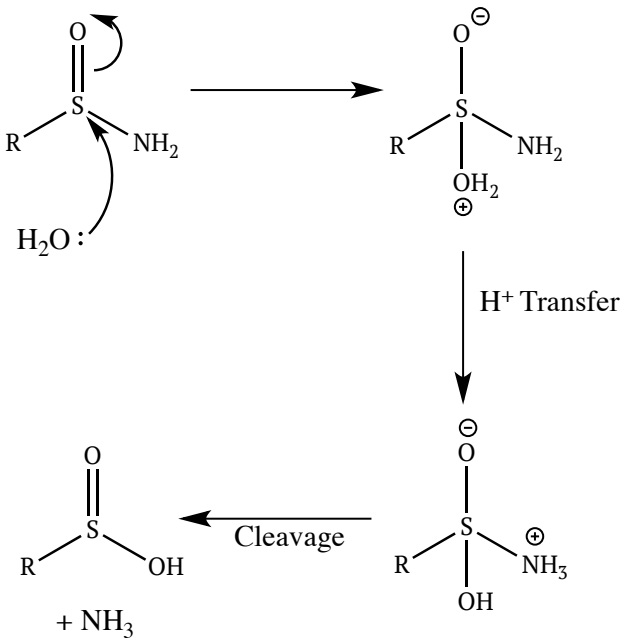
Asefi and Tillet reported the first study of acid-catalysed hydrolysis of *p*-tolyltoluene-*p*-sulfinamide and determined that water was acting as the proton transfer agent.[83] This work also makes the first reference to the potential sites of protonation in the sulfinamide molecule, citing both the oxygen and nitrogen atoms as possibilities. However, no conclusion was made as to the more likely site for protonation of the sulfinamide group. An overall scheme for acid

catalysed hydrolysis of sulfinamides in the presence of hydrochloric acid was proposed (Scheme 11).



Scheme 11: Acid (HCl) catalysed hydrolysis of sulfinamide 31Q [adapted from reference].[83]

Investigations into the impact of molecular structure on sulfinamide hydrolysis were completed in D_2O by Wagner *et al.*[84] The work reported that acyclic sulfinamides hydrolyse far more rapidly than their cyclic counterparts, due to rapid rotation along the S(O)-N bond; five-membered ring sulfinamides hydrolyse faster than six-membered ring sulfinamides; and that a five membered ring sulfinamide with a proton on the nitrogen hydrolysed twice as fast as a five membered ring sulfinamide with an alkyl group on the nitrogen. Consideration was given to the mechanism here, with water solvation playing a significant role. The study suggested that water attacks the sulfur first, followed by proton transfer to the nitrogen, leading to charge separation along the S-N bond and subsequent cleavage of the zwitterion (Scheme 12).



Scheme 12: Mechanism for sulfinamide hydrolysis proposed by Wagner *et al.*[84]

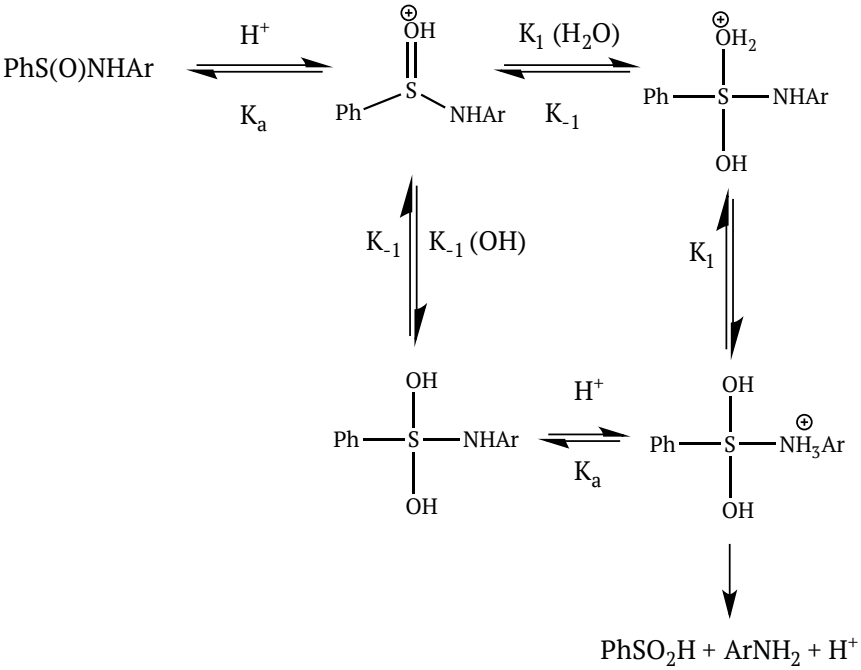
More directed work into the site of protonation was conducted by Bagno *et al*., and published in 1994. ^{14}N NMR relaxation studies lead to the conclusion that the site of protonation on the sulfinamide was the oxygen atom, with only a slight broadening seen in the ^{14}N spectrum during hydrolysis.[85] In contrast, Bujnicki *et al.* studied both the free and protonated forms of sulfinamides using IR and ^{14}N NMR, concluding that protonation of sulfinamides occurs on the nitrogen.[86] However, they did remark that their studies could not exclude some degree of protonation on the oxygen atom. *Ab initio* calculations on gas phase acid hydrolysis of *N*-methylmethane sulfinamide performed by Kim and Lee in 1997 remarked that the *O*-protonated form of the sulfinamide was more stable by 10.6 kcal mol^{-1}.[87]

Kim and Lee[87] also investigated the effect of substitution on the nitrogen atom. The work determined that as the substituent is varied, there is no change in the activation energy for proton transfer. The study made four clear conclusions:

1. The first step of hydrolysis is protonation, occurring on the oxygen atom.
2. Addition of water takes place to form a sulfurane intermediate, which breaks via S-N bond cleavage to form products.
3. 1,3 hydrogen transfer occurs to the nitrogen in a 'barrierless' process to yield the products, which are the sulfinic acid and quaternary ammonium NHR_3^+.

4. The rate determining step is the breakdown of the protonated intermediate.

The work of Kim and Lee built upon the ^{18}O exchange experimentation in perchloric acid conducted by Okuyama *et al*. in 1994.[88] The pH rate profile observed was consistent with a two-step mechanism *via* the hypervalent sulfurane reaction intermediate shown in Scheme 13. ^{18}O exchange experiments confirmed that the rate-determining step for hydrolysis is the breakdown of this sulfurane intermediate. Oxygen exchange was seen in unreacted starting sulfinamide, which could only occur via the proposed sulfurane intermediate.

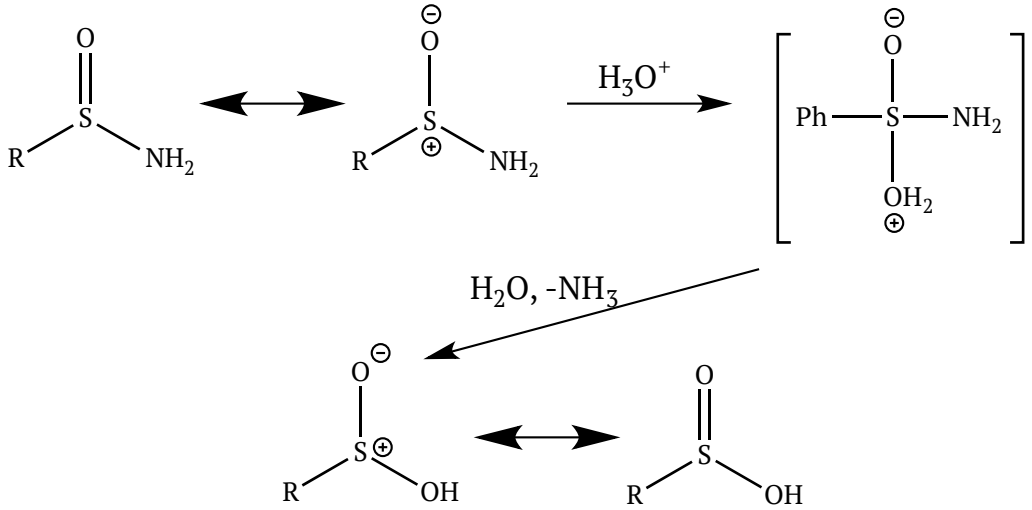


Scheme 13: Hydrolysis mechanism *via* the hypervalent sulfurane intermediate.

Mechanistic investigations into the general acid catalysis of secondary sulfinamides have been performed more recently (2007) by Piggott and Karuso using *pseudo* first order conditions with a phosphate buffer.[89] The work suggested that the reaction can be catalysed by a general acid/base species, but regardless of the site of protonation (N or O), the protonation event is not the rate-limiting step. It was noted that the presence of an aromatic ring on the sulfur only slightly increases the rate of hydrolysis, but that the rate of hydrolysis is fastest for tertiary sulfinamides. The slower rate of hydrolysis of secondary sulfinamides was attributed to stabilisation of the S=O bond by strong intermolecular hydrogen bonding to the N-H. Extrapolating from this, it might

suggest that primary sulfinamides are even less susceptible to hydrolysis, since they can participate in more hydrogen bonding interactions than their secondary counterparts, although this was not commented on in their work.

In 2013, a study of hydrolysis using primary sulfinamides was published with a view to understanding peptide sulfinamide hydrolysis reactions at physiological pH.[90] Ammonia was confirmed as the only other by-product of the hydrolysis reactions of primary sulfinamides, and the results were consistent with the agreed mechanism *via* the sulfurane intermediate (Scheme 14).



Scheme 14: Proposed mechanism of hydrolysis of primary sulfinamides (reproduced from Toscano *et al.*).[10]

The literature evidence to date has not yet reached a consensus as to the exact mechanistic pathway *via* which the sulfinic acid product is formed; the existence of the sulfurane intermediary has become an accepted concept. However, both sites of protonation (O or N) are individually substantiated by research. The majority of the research available relates to secondary and tertiary sulfinamides, along with acid/base conditions and so cannot be used as a direct comparison for the reaction conditions that have resulted in the products prepared in this project.

To investigate the existence of a trend in the results obtained for this series of materials, the crystallization conditions were examined in detail. The salt products were obtained in bulk methanol, and ethanol, in addition to aprotic acetonitrile, and therefore solvent mediated alcoholysis as described in 2015[91] was excluded as a potential reaction pathway by which the hydrolysis could

was excluded as a potential reaction pathway by which the hydrolysis could occur. This 2015 study, performed by Mikołajczyk and coworkers,[91] used acid catalysis with perfluoroacetic acid; entirely different to the conditions used in this work. As the conditions used in our study were examined more closely, atmospheric oxygen seemed to play a key role in determining the formation of the sulfinate versus the sulfonate salt.

Where the crystal structure determination was performed in quick succession following the evaporation of the crystallization solvent to dryness, the sulfinate salt was recovered (**55A**, **55C** and **55I**). In the cases of **56A**, **56L**, **56M**, and **56O**, a significant amount of time had passed between the initial crystallization experiments and the structural analysis, and the sulfonate salt was recovered in each case. For the unsubstituted phenyl analogue (**55A** and **56A**), both of the salts could be observed. The sulfinate salt was formed under dry conditions in a dessicator, despite attempts to remove all traces of water; and the sulfonate salt formed through normal crystallization on the bench from methanol.

The ease of hydrolysis is related to the dihedral angle of the sulfinamide S=O relative to the aryl plane, calculated as shown in Figure 43 and shown in Table 10 (calculated by Mercury structure visualisation software).[45]

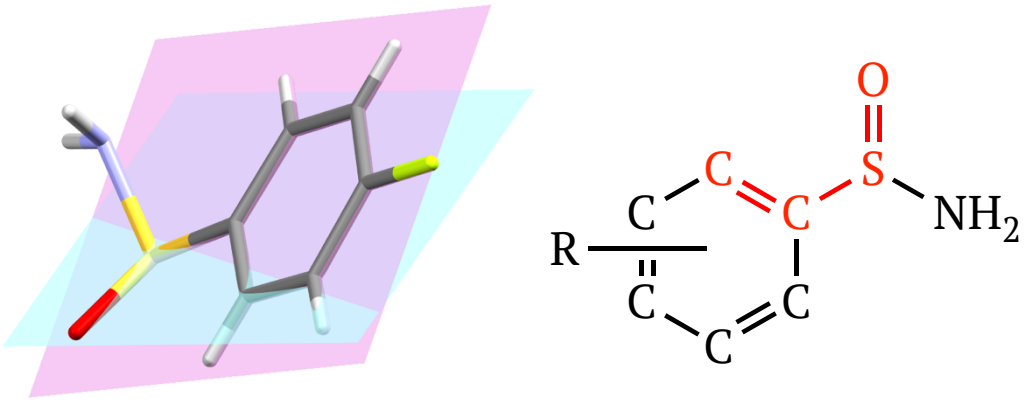


Figure 43: The relative dihedral angle for sulfinamides S=O relative to the aryl ring (left) as calculated through the O-S-C-C bonds (right, red).[45]

Table 10: The dihedral angle in the sulfinamide crystal structures. (structures for which sulfinate/sulfonate salt was also recovered are highlighted in green).

Sulfinamide	Dihedral Angle
2-Br [**31C**]	7.10°
3-Br [**31D**]	25.3°
4-Br [**31E**][a]	24.4°
4-Br [**31E**][a]	39.9°
3-Cl [**31G**]	-26.2°
3-F [**31J**]	26.1°
4-F [**31K**]	28.1°
4-Me [**31Q**][1]	24.3°
2-MeO [**31L**]	16.1°
3-MeO [**31M**]	-1.1°
4-MeO [**31N**]	-28.1°
3-NO_2 [**31S**]	-10.2°
4-NO_2 [**31T**]	-26.2°

a - Two dihedral angles occur in this crystal structure as Z' = 2;

Thus, when the S=O bond is close to co-planar with the aryl ring, hydrolysis is observed. This can be interpreted in a number of ways. On a steric basis, access to the S=O bond may be physically easier in the co-planar conformation, allowing for more efficient attack of water on the sulfinamide; or alternatively, conjugation of the S=O orbitals and the aryl ring alters the electronic susceptibility of the S=O moiety to protonation and nucleophilic attack by water.

A small dihedral angle should suggest a better level of conjugation between the S=O bond and the aromatic ring, thereby leading to the S=O bond having greater double bond character and reducing electron density on the oxygen atom. In such instances, the sulfinamide nitrogen would possess the greater electron density, thereby leading to preferential protonation of the nitrogen atom. The resultant protonated sulfinamide then would possess a good leaving group (ammonia) which can be eliminated after nucleophilic attack on the sulfur by adventitious water present in the solvent. Without this additional conjugation, protonation of the oxygen may be thought to occur either along with, or preferentially to, protonation of nitrogen. In this case, elimination of ammonia, subsequent to nucleophilic attack by water, is likely to be slower as the additional proton transfer step (as described by Kim and Lee)[87] is required to create the neutral leaving group.

After formation of the hydrolysis product, the sulfinic acid reacts with the ammonia produced during the hydrolysis reaction to form the corresponding ammonium salt. Where the crystallizations were performed under a nitrogen atmosphere, or analysed in quick succession after crystallization, the ammonium sulfinate materials were isolated and characterised. It is most likely that, if any hydrolysed material were left open to air for a sufficient time, the material would further react with oxygen to form the ammonium sulfonate salt.

The unsubstituted benzene sulfinate and sulfonate salts (**55A** and **56A**) were isolated from differing reaction conditions (Table 9). However, it required several single solvent crystallizations of the parent sulfinamide to isolate any crystals suitable for structure solution, and the parent sulfinamide structure could not be determined at all. Since the sulfinate/sulfonate salts diffracted very well in all cases, it could be concluded that the parent sulfinamide was slightly more stable with respect to hydrolysis, but poorly crystalline overall, and so structure determination of the pure sulfinamide was not possible.

It was interesting to note that the majority of the hydrolysed structures were 2-substituted. *Ortho*-substitution could affect the dihedral angle by forcing the S=O bond to orient itself away from the substituent, and towards co-planarity with the aryl ring. This was observed in **31C** (2-Br) and **31L** (2-MeO), creating a very small dihedral angle, and potentially, the optimum situation for hydrolysis.

A second point to note is that structure determination was much easier where the sulfinamides contained electron withdrawing substituents (halo or nitro), and these materials appeared to be more stable with respect to hydrolysis than those with electron donating substituents (methoxy or methyl). Further investigation would be required to fully understand how electronic effects influence these materials.

Interestingly, the dihedral angle observed for the 2-methoxy analogue [**31L**] was determined to be 16.1°. This is larger than that seen for other materials that exhibited hydrolysis; and also larger than that observed for the 3-nitro analogue (-10.2°) which seemed to be stable under the crystallization conditions. It is

possible that the influence of the electron withdrawing nitro substituent is contributing to a higher level of stability in the case of the 3-nitro material.

It should be noted that **31L** did appear to be slightly more stable with respect to crystallization conditions when compared to the other materials for which hydrolysis was observed. The crystal structure of **31L** was obtained from bulk solvent, but only where crystallization was done quickly on a warm clock glass and SCXRD analysis performed almost immediately, otherwise the hydrolysed material was crystallized. As discussed earlier, **31L** lends itself to the ingress of water molecules into the channels, which may facilitate the hydrolysis of the compound, more so than the influence of the dihedral angle.

The 3-methoxy analogue [**31M**] also forms channels suitable for water ingress similar to **31L** and exists in the more co-planar conformation [relative angle 1.1°], associated with increased rate of hydrolysis. This material seemed to be extremely unstable with respect to hydrolysis, indicating that these factors are both partly contributing to the overall reactivity of the molecule.

Further insight into the influence of the substituent on the dihedral angles around the S=O bond can be inferred if the crystal structures of the sulfinate/sulfonate salts are compared. It seems that where the substituent is in the ortho position, this forces one of the S=O bonds into the plane of the aromatic ring (Table 11). In contrast, the unsubstituted benzene sulfinate and sulfonate S=O bonds orient themselves out of the plane of the aryl ring. The trend correlates well with the connection made between the ease of hydrolysis and the dihedral angle for the parent sulfinamide materials (above, Table 10).

Table 11: Dihedral angles in Sulfinate and Sulfonate Salts [S=O forced into plane of ring highlighted in blue].

Material		S=O dihedral angle 1	S=O dihedral angle 2	S=O dihedral angle 3	Sulfinamide S=O dihedral angle
Sulfinate Salt	**55A** (R=H)	38.7°	152.3°	-	-
	55C (2-Br)	6.6°	118.6°	-	7.1°
	55I (2-F)	0.8°	111.7°	-	-
		2.1°	111.5°	-	-

Sulfonate Salt	**56A** (R=H)	18.0°	100.5°	141.6°	-
	56F (2-Cl)	-0.4°	121.1°	118.5°	-
	56L (2-MeO)	-3.4°	124.5°	116.0°	16.1°
		-5.7°	125.5°	113.5°	16.1°
	56M[a] (3-MeO)	-0.94°	122.3°	119.2°	-1.1°
	56O (2-Me)	5.6°	124.8°	113.6°	-

a – Calculated on the side of the methoxy substituent as material is 3-substituted, for all others angle was calculated relative to the first unsubstituted *ortho*-hydrogen atom on the aromatic ring;

As mentioned previously, the first crystal structure of this type [**56F**] was determined from crystals found in the acetonitrile solution used in mass spectrometry analysis. The positive and negative mode comparison in the nominal mass spectra provided an interesting insight into the progression of the materials from sulfinamide through to the sulfonate salt (Figure 44).

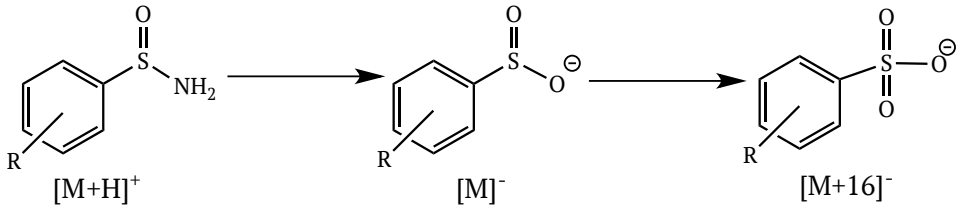


Figure 44: Mass spectrum progression of sulfinamide to sulfonate.

The mass spectrum clearly showed the molecular change from the parent sulfinamide [$(M+H)^+$] in the positive mode, to the sulfinate salt [$(M)^-$] in negative mode, in addition to formation of the sulfonate salt [$(M+16)^-$] in negative mode also. However, it should be noted that the hydrolysis could be catalysed under mass spectrum conditions by the small amount of formic acid used in the analysis. Both sulfinate and sulfonate were observed in the mass spectra for all of the 18 sulfinamides in the series, suggesting that the correct conditions would lead to hydrolysis for all of the sulfinamides.

2.5.2 Crystal Structures of Sulfinate and Sulfonate Salts

The CSD[2] contains crystal structures of 32 aromatic ammonium sulfonate salts, 7 of which were simple mono-substituted aryl salts, similar to those determined in this study.[92–100] This group includes ammonium 4-nitrobenzene sulfonate (**56T**,

VEPSIU),[97] ammonium 4-methoxybenzene sulfonate (**56N**, XEBQUS)[98] and three polymorphic forms of ammonium 4-methylbenzene sulfonate (**56Q**, DUTZEX01/02/03),[93,95,99] which are the sulfonate analogues of **31N**, **31Q** and **31T** respectively.

The structures of these three materials are remarkably similar to each other, and to those determined in this work, displaying a repetitive pattern of strong N-H···O=S hydrogen bonds extending in a linear fashion down the unit cell axes, as shown for the 4-methoxy substitituted material (**56N**, XEBQUS, Figure 45).[98] Interestingly, there are no hydrogen bonds formed to the methoxy groups in this structure, likely due to the orientation of the molecules required to support the stronger hydrogen bonding interactions to the S=O bond.

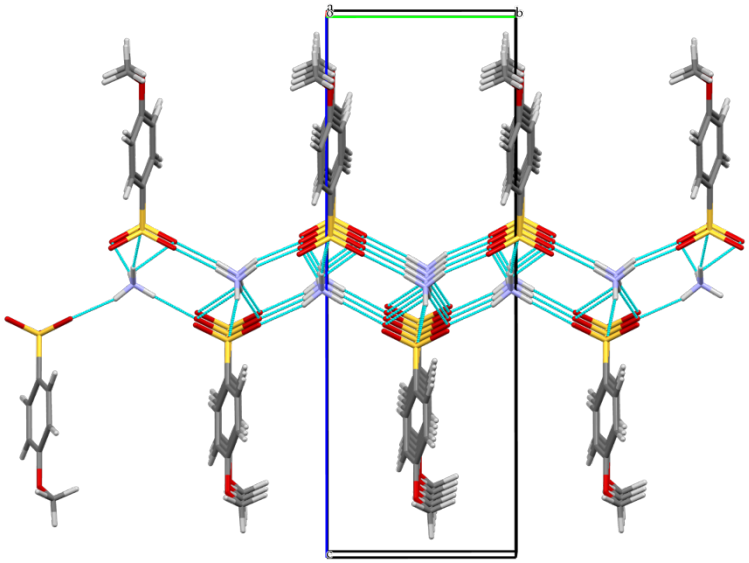


Figure 45: Linear extension of hydrogen bonds along the *b*-axis and down the *a*-axis in XEBQUS [4-MeO, 56N].

The 4-nitro substituted analogue (**56T**, VEPSIU) displays rotational disorder around the sulfonate anion, but it is possible to observe the characteristic interactions none the less (Figure 46).

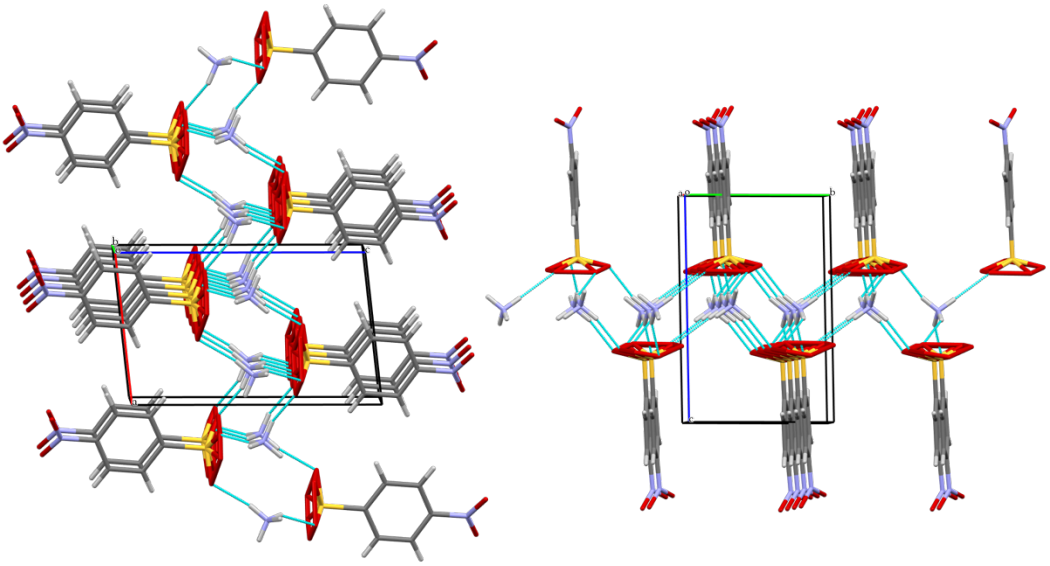


Figure 46: Linear extension of hydrogen bonds in VEPSIU (4-NO_2, 56T).[97]

These strongly defined interactions result in the formation of hydrophilic and hydrophobic areas within all three structures, as can be observed in the example of ammonium 4-methylbenzene sulfonate (**56Q**, DUTZEX01, Figure 47).[93]

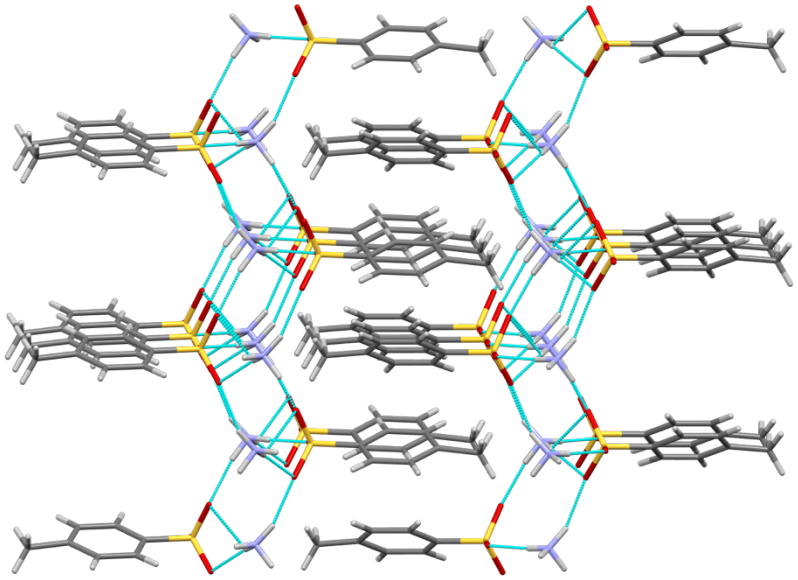


Figure 47: hydrophilic and hydrophobic areas in DUTZEX01 [4-Me, 56Q].[93]

The sulfinate and sulfonate salts determined in this study display linear extension of hydrogen bonding along two unit cell axes, in conjunction with strong hydrogen bonds to the water molecules in the case of **55A** and **55C** (Table 12). The overall supramolecular architecture in these materials forms similar hydrophilic and hydrophobic areas within the structures.

Table 12: Hydrogen bonding interaction observed in ammonium sulfinate/sulfonate salts.

Compound Number		Space Group	C(4) directionality	Hydrophilic and hydrophobic planes
Sulfinate Salt	**55A**[a] (R = H)	*Pbca*	*a*-axis *b*-axis	*ac*-plane *bc*-plane
	55C[a] (2-Br)	$P2_1/c$	*b*-axis *c*-axis	*ab*-plane *ac*-plane
	55I (2-F)	$P2_1$	*a*-axis *b*-axis	*ac*-plane *bc*-plane
Sulfonate Salt	**56A** (R = H)	*Pbca*	*a*-axis *b*-axis	*ac*-plane *bc*-plane
	56F (2-Cl)	$P2_1/c$	*b*-axis *c*-axis	*ab*-plane *ac*-plane
	56L (2-MeO)	*Pc*	*b*-axis *c*-axis	*ab*-plane *ac*-plane
	56M (3-MeO)	$P2_1/c$	*b*-axis *c*-axis	*ab*-plane *ac*-plane
	56O (2-Me)	$P2_1/c$	*b*-axis *c*-axis	*ab*-plane *ac*-plane
	56N (4-MeO) XEBQUS[98]	$P2_12_12_1$	*a*-axis *b*-axis	*ac*-plane *bc*-plane
	56Q (4-Me) DUTZEX01[93]	$Pn2_1a$	*b*-axis *c*-axis	*ab*-plane *ac*-plane
	56T (4-NO_2) VESPIU[97]	*P*-1	*a*-axis *b*-axis	*ac*-plane *bc*-plane

a – Structure determined as a hydrate;

The structures of **55A** and **55C** and **55I** represent the first ammonium sulfinate salt crystal structures reported to date. The crystal structures found in the CSD[2] and those determined in this study exhibit the same expected pattern of strong N-H···O=S hydrogen bonds (1.9-2.2 Å) from the S=O anion to the ammonium functional group.

Two of the three sulfinate structures were determined as hydrates, with the water molecules hydrogen bonded to the S=O of the sulfinate molecules, forming chains. The growth in unsubstituted **55A** along the *a*- and *b*-axes contributes to

the formation of hydrophilic and hydrophobic segregation in the *ac*- and *bc*-planes (Figure 48, Figure 49 and Figure 50).

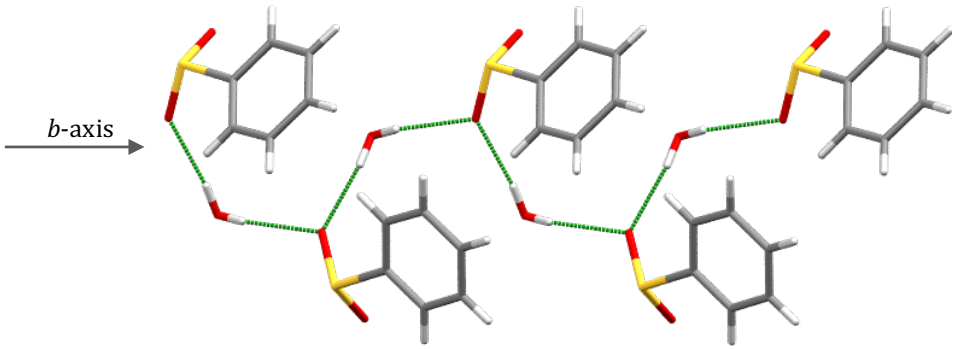


Figure 48: Interactions with water molecules observed in 55A (R=H).

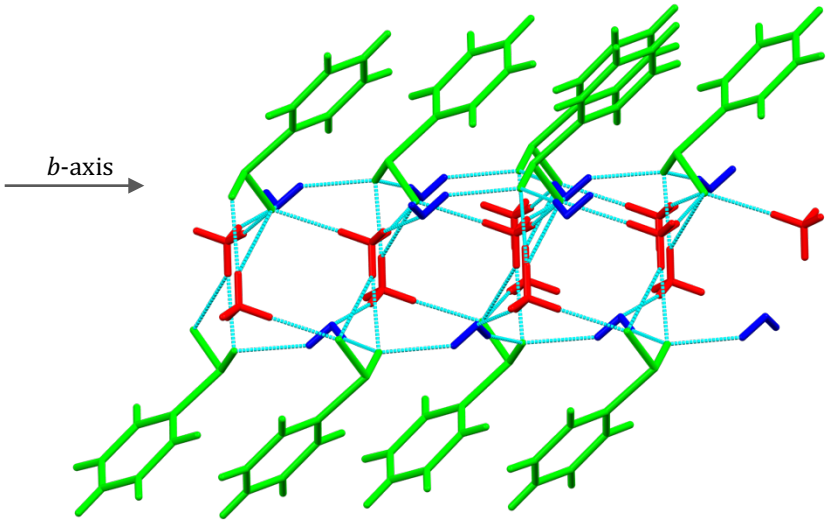


Figure 49: Interactions along unit cell axes, with water molecules (blue) and ammonium ions (red) as observed in the crystal structure of 55A (R=H).

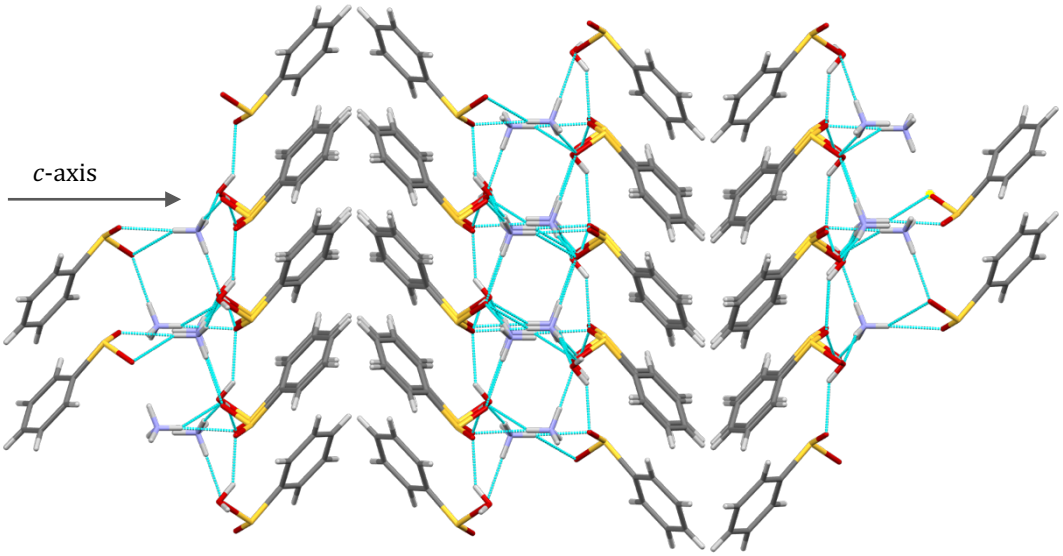


Figure 50: Formation of hydrophilic and hydrophobic areas in 55A.

55C (2-Br) displays the same hydrophilic and hydrophobic segregation, with the chains to ammonium ions contributing to formation of chains spiralling down the *c*-axis *via* hydrogen bonding with the water molecules (Figure 51).

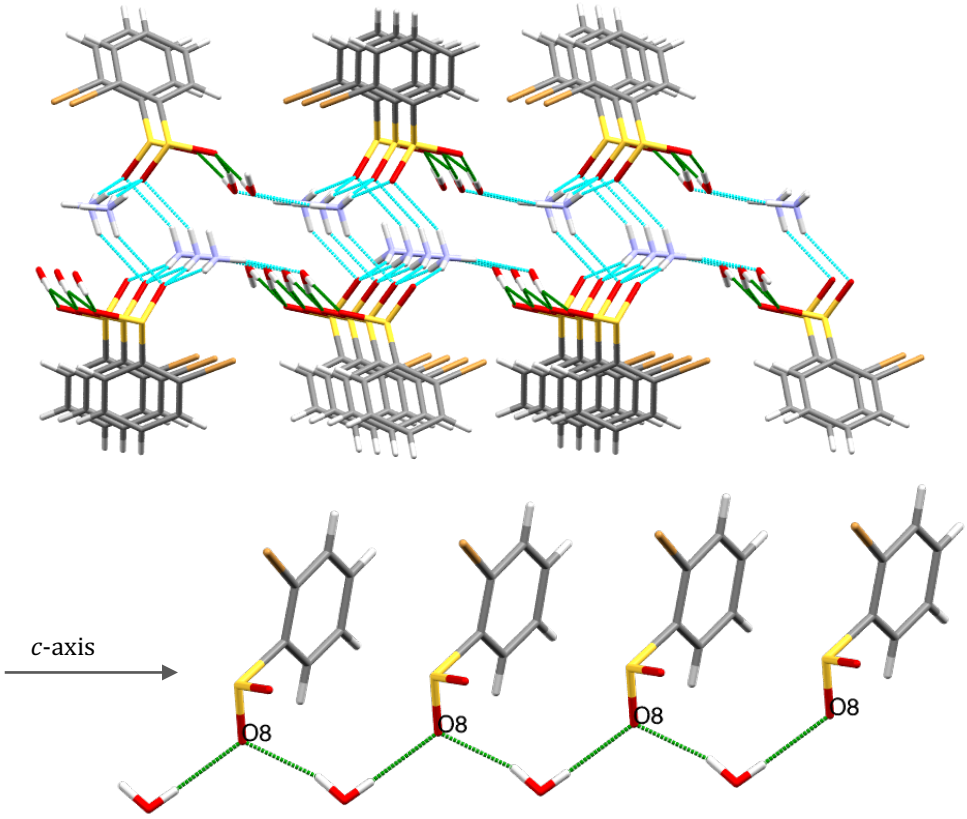


Figure 51: Chains in 55C forming spirals as viewed down the *c*-axis (top) and along the *c*-axis (bottom).

55I (2-F), which is anhydrous, displays the same hydrogen bonds from the S=O anion to the ammonium cation, with alternating orientations of the aromatic ring facilitating these interactions (Figure 52).

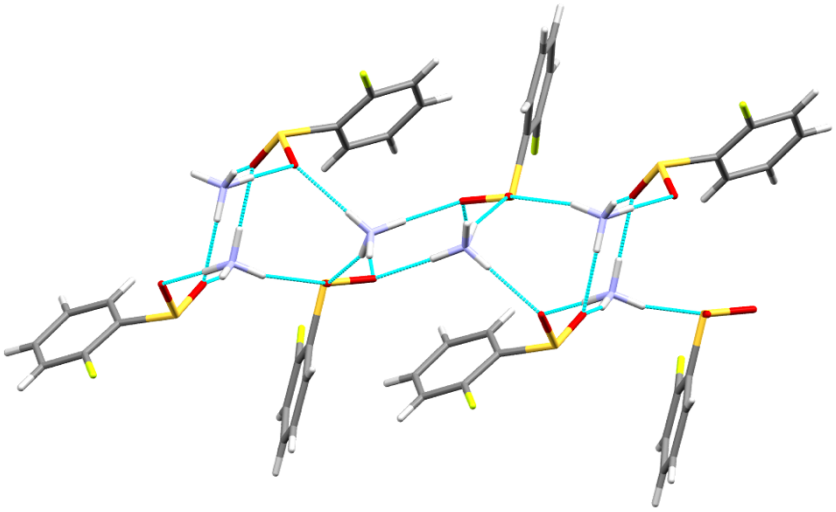


Figure 52: Alternating orientation of aromatic rings in 55I (view down the *c*–axis).

The only instance where the sulfinate and sulfonate structures can be compared directly is in the case of the unsubstituted derivatives **55A** and **56A**. As seen in Table 12, **55A** and **56A** crystallize in the same space group, with similar hydrophilic and hydrophobic regions in the structures. In comparison to its sulfinate counterpart [**55A**] the aromatic rings in the sulfonate derivative **56A** are almost perpendicular relative to each other (82°), presumably in order to facilitate interactions to the additional strong hydrogen bond acceptor (S=O) in the sulfonate moiety (Figure 53). There is also an additional C-H···O=S bond from the aromatic ring to the nearest sulfonate oxygen in **56A** [2.4 Å] that is not observed in the sulfinate counterpart, this might account for the slight difference in angles between the aromatic rings (Figure 54). The corresponding sulfinate structure (above) has a relative angle of 45° between the aromatic rings. Apart from these slight variations, there is little to differentiate between the structures, even though **55A** is hydrated and **56A** is anhydrous (Figure 55).

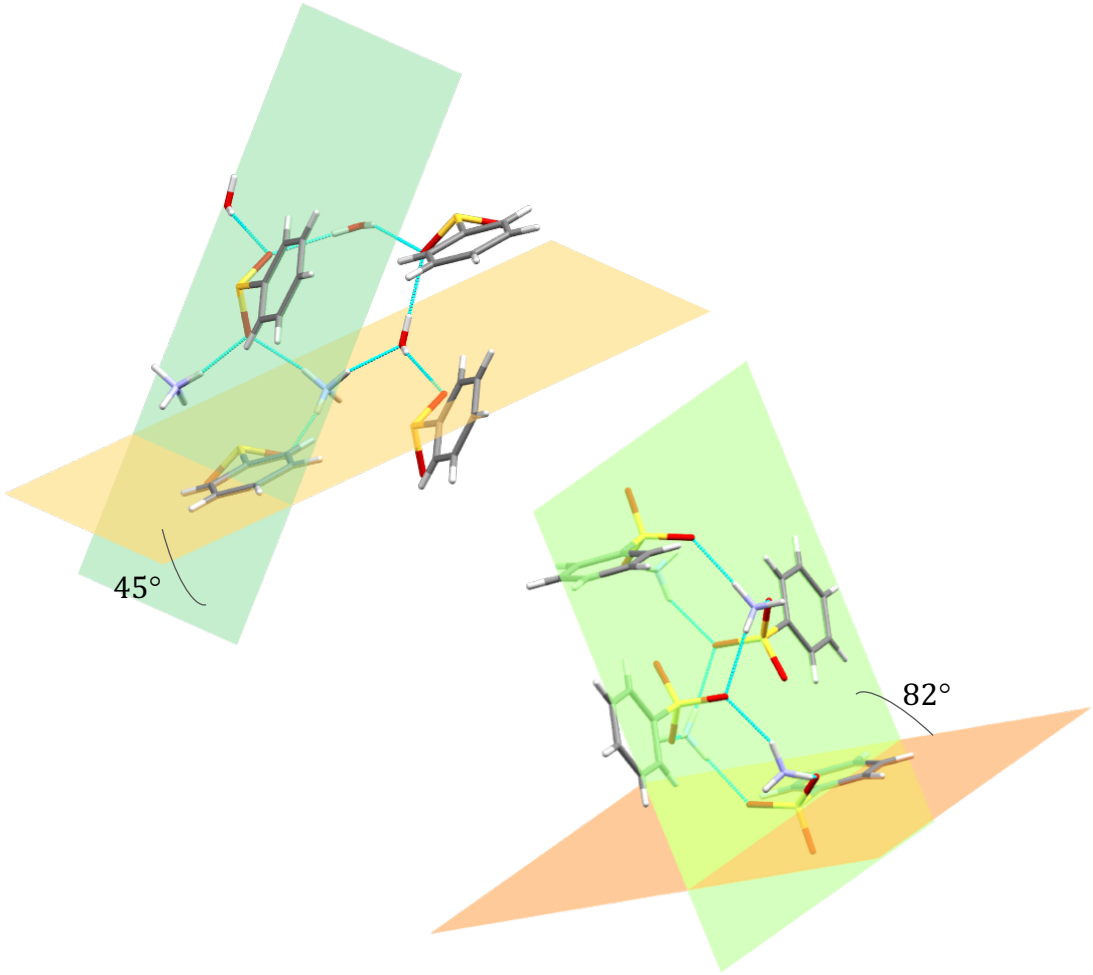


Figure 53: Angularity of aromatic rings in 55A (top) and 56A (bottom) [R=H].

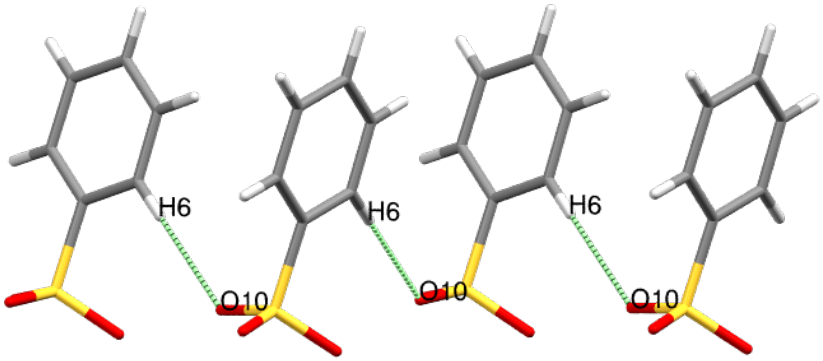


Figure 54: C-H···O=S bond in 56A.

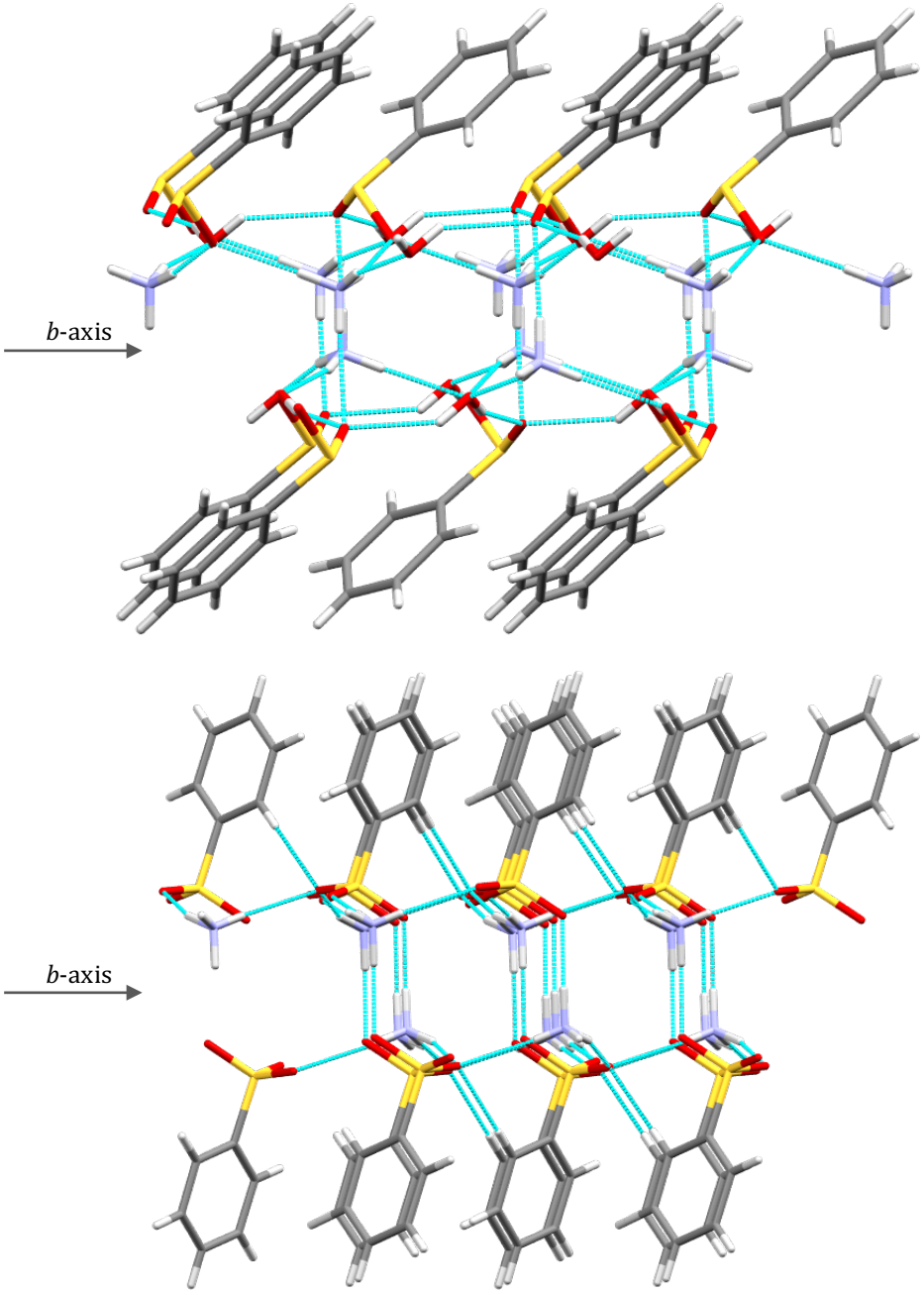


Figure 55: Similarities in the structures of 55A (top) and 56A (bottom).

The sulfonate structures of the 2-methoxy and 3-methoxy derivatives (**56L** and **56M**) display a similar system of intermolecular N-H···O=S hydrogen bonds, creating strongly defined hydrophilic and hydrophobic segregation within the crystal structures with the aromatic rings almost perpendicular to each other [approx. 85°] (Figure 56 and Figure 57).

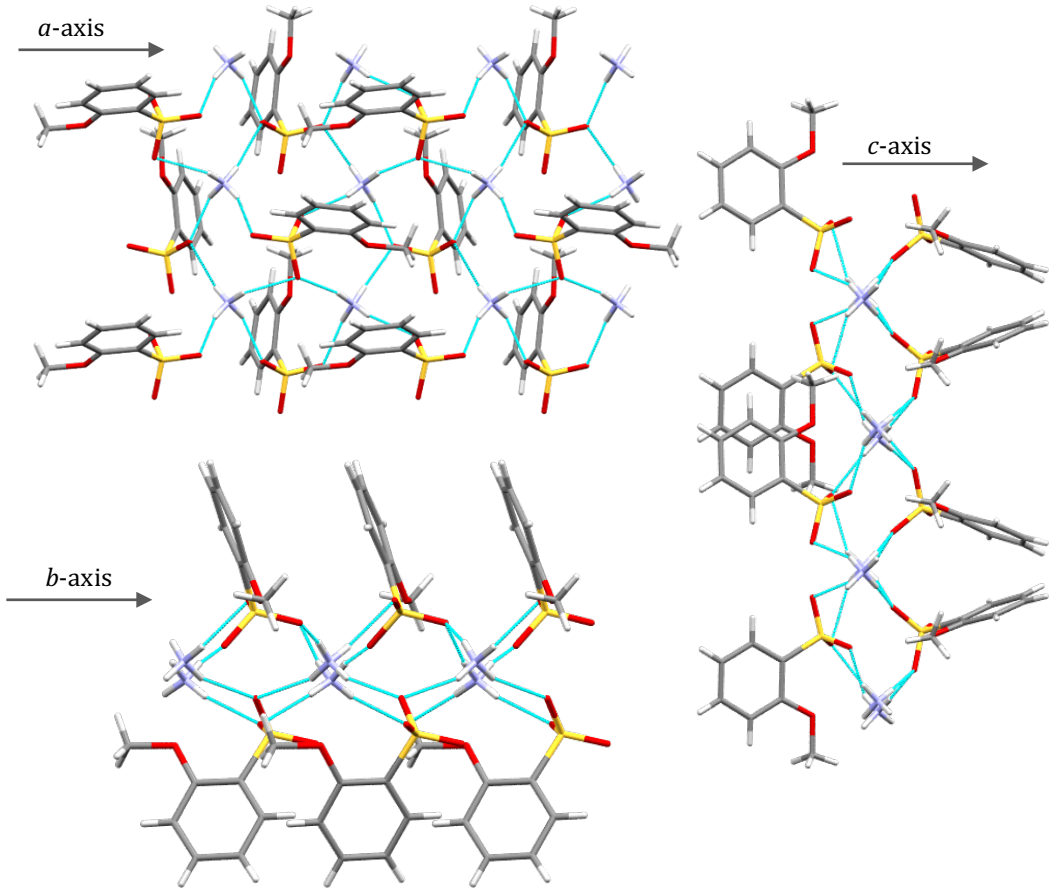


Figure 56: Interactions along the unit cell axes as observed in 56M (2-MeO).

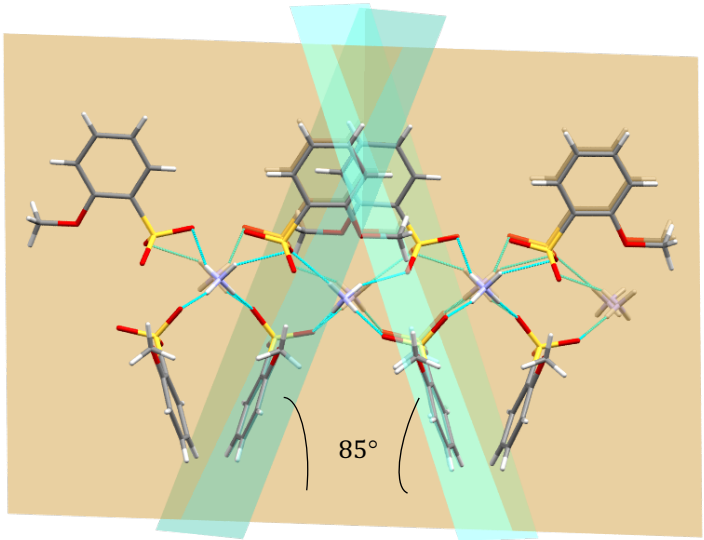


Figure 57: Relative angles of aromatic rings in 56L (2-MeO).

Additional interactions are observed as short contacts along with those mentioned above, although the validity of these as true hydrogen bonds is debatable. **56L** displays an apparent N-H···O-C interaction from the ammonium cation to the methoxy group (Figure 58, approx. 2.8 Å). Since these interactions do not contribute to a change in the overall architecture of the crystal structure, they could again be considered as structural artefacts, with the methoxy oxygen only accepting a hydrogen bond opportunistically due to the high availability of donors in the vicinity. This is further substantiated by the fact that the 4-methoxy analogue (**56N**, XEBQUS)[97] displays no interactions of this type.

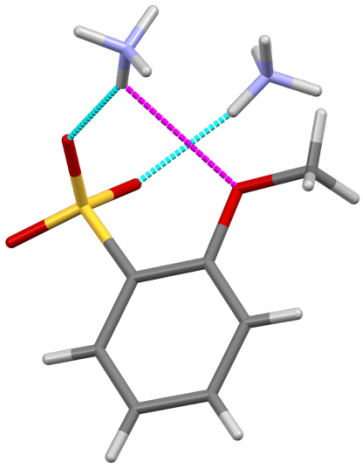


Figure 58: N-H···O-C interaction observed in 56L [2-MeO] (magenta and cyan).

56M (3-MeO) displays an $R_2^2(4)$ tetramer formed within the hydrophilic area of the ammonium to sulfonate interactions (Figure 59). This structure also contains the same $R_2^2(6)$ dimer linking the methoxy groups (Figure 60) that was observed previously for **31N**, 4-methoxybenzene sulfinamide. The orientation of the aromatic rings in this crystal structure appears to be unaffected when compared to the other structures determined in this series (Figure 61), and the dimer is not contained in the crystal structure of **56L** or **56N** (XEBQUS),[98] and so it is likely that this dimer is also a structural artefact, rather than a defining feature.

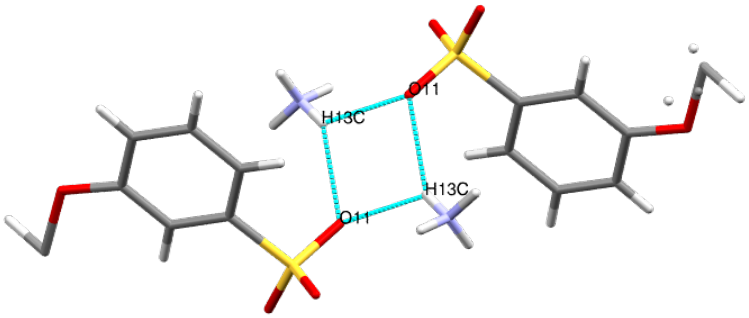


Figure 59: $R_2^2(4)$ tetramer formed in 56M (3-MeO) [the CH_3 group is disordered].

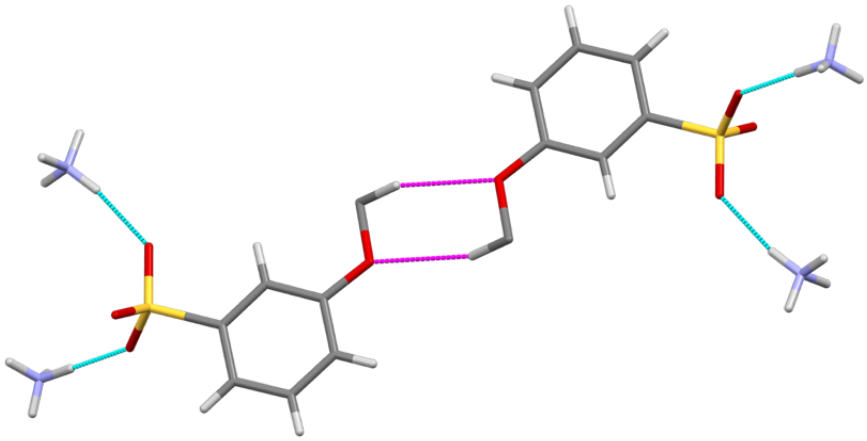


Figure 60: $R_2^2(6)$ dimer observed in 56M (magenta), [the CH_3 group is disordered].

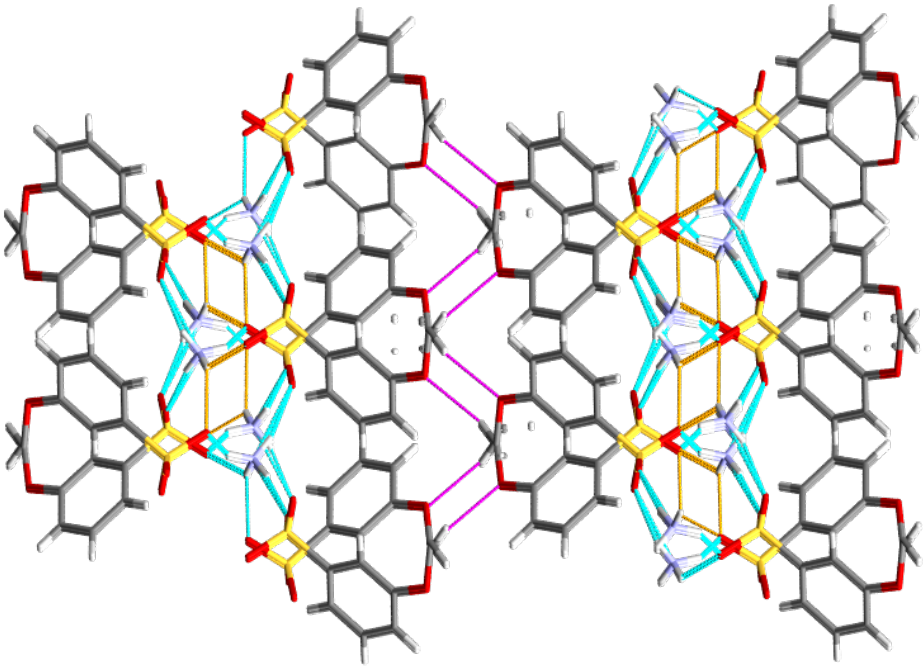


Figure 61: Overall packing including $R_2^2(4)$ tetramer (orange) and $R_2^2(6)$ dimer (magenta) observed in 56M (ammonium hydrogen bonds in cyan) [the CH_3 group is disordered].

The 2-chloro derivative [**56F**] displays the anticipated system of hydrophilic and hydrophobic hydrogen bonds, with an almost wave-like pattern of the hydrogen bonds as they spiral down the *c*-axis. The chlorine atom is held at alternating positions in the hydrophobic region of the layers of molecules (Figure 62).

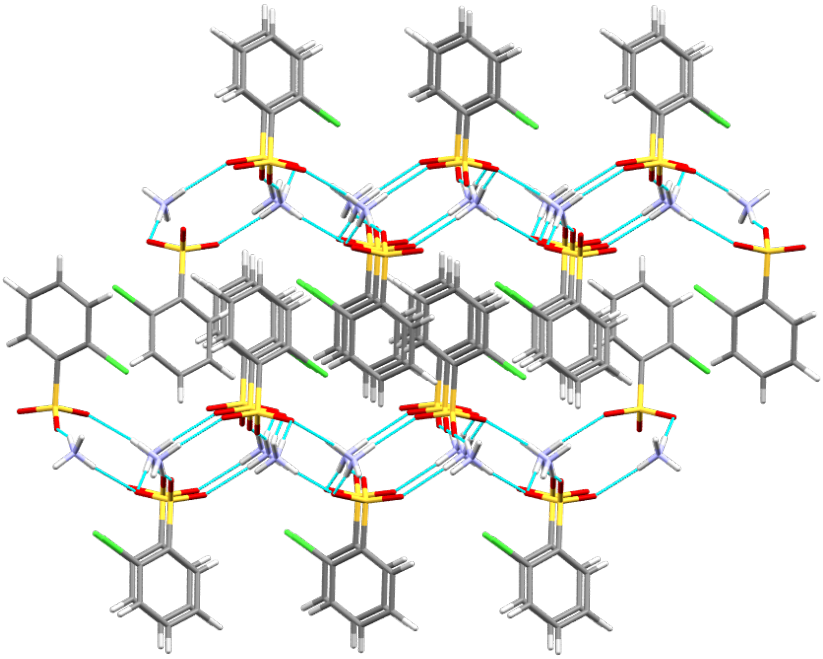


Figure 62: Wave-like hydrogen bonding interactions in 56F (2-Cl).

The final structure, 2-methyl benzene sulfonate [**56O**], displays an $R_4^4(12)$ tetramer formed between two sulfonate and two ammonium cations (Figure 63) and a similar pattern of hydrogen bonds to the preceding structures.

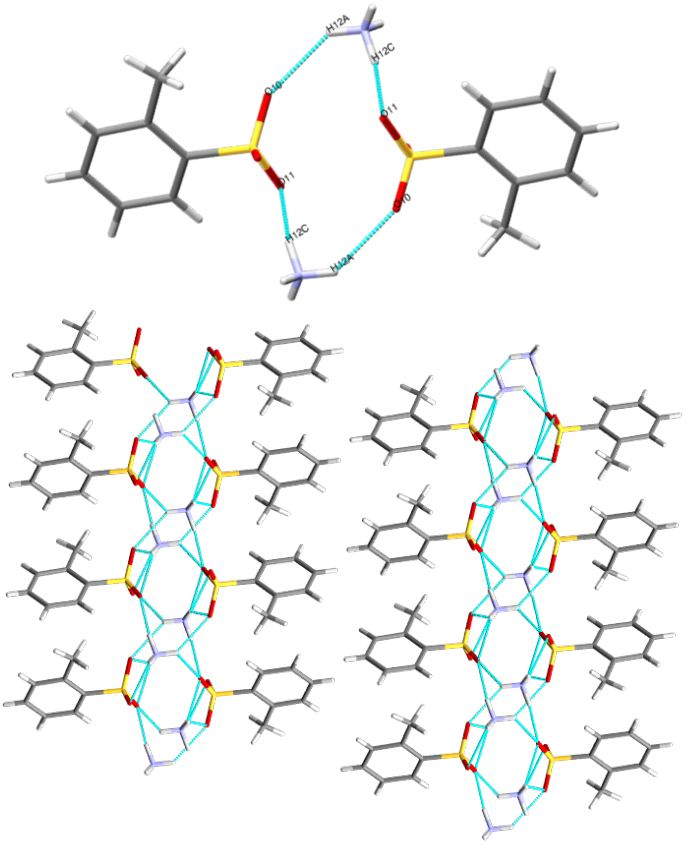


Figure 63: The $R_4^4(12)$ tetramer formed within the crystal structure of 56O (2-Me).

Overall, the sulfinate and sulfonate salts display the anticipated pattern of strong hydrogen bonds from the sulfinate/sulfonate S=O to the ammonium ions, with the participation of water within the hydrated crystal structures having little effect upon the overall supramolecular architecture. The overall crystal packing in all structures is distinguished by well-defined hydrophilic and hydrophobic areas and linear crystalline extension along the unit cell axes, and correlates very well with the literature examples available for comparison.

2.5.3 Analytical Characterisation of Sulfinate and Sulfonate Salts

In order to fully characterise these novel materials, as well as ascertain whether the determined crystal structure was representative of the bulk material, efforts were directed toward preparation of pure samples of the sulfinate/sulfonate salts by dissolution of the parent sulfinamide in aqueous ethanol at room temperature open to the air, and allowing the solution to stir for a period of 1 week before checking for reaction completion. Reaction completion was

determined by the appearance of one set of signals in ^{1}H NMR. Conversion to product was observed after 1 week in the cases of **55A**, **55I** and **56F**. Full characterisation of these materials was successfully performed with removal of the water under reduced pressure with ether as an azeotrope, followed by analysis via NMR in $CDCl_3$. Evidence of molecular change was evident in the NMR spectrum (as discussed previously).

The 2-bromo analogue [**55C**] was incompletely converted after one week of stirring in aqueous ethanol. NMR analysis of the 1 week sample determined that 50% of **31C** had converted to the product (Figure 64). The comparison of the NMR spectra in $CDCl_3$ shows peaks remaining from the sulfinamide starting material, and those which can be attributed to the hydrolysed product (including the ammonium at approx. 7.25 ppm); the proton signals can be integrated equally (four aromatic protons for each molecule, the 4 ammonium protons and the 2 sulfinamide N-H protons), indicating 50% conversion to product. In order to fully convert this sample to product, the solution was stirred for another two weeks.

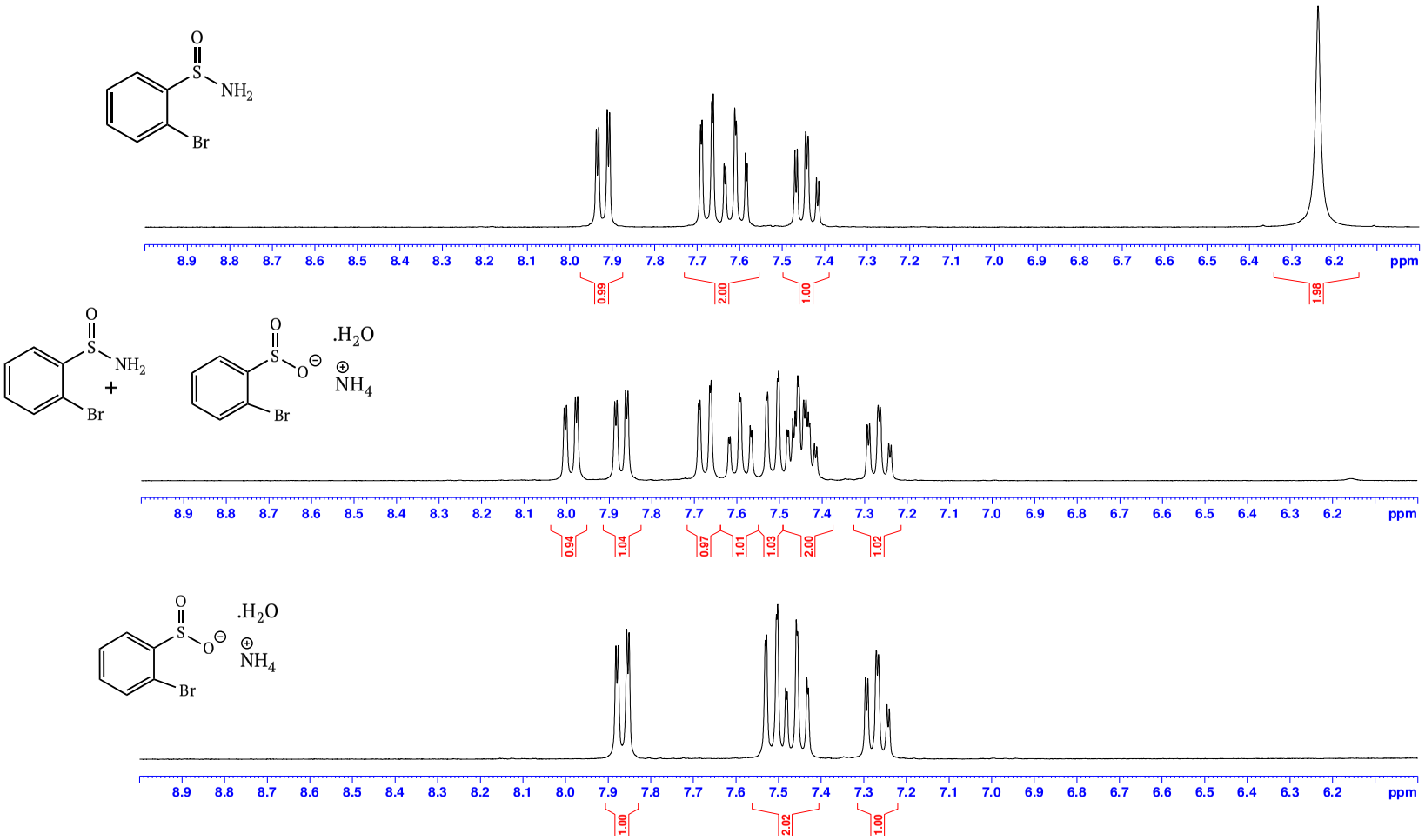


Figure 64: NMR comparison of 31C [2-Br] (DMSO-d_6, top), product after 1 week stir in aqueous ethanol (MeOD, centre) and fully converted to 55C after 3 weeks (MeOD, bottom).

In the case of **56L**, stirring in aqueous ethanol caused structural change in the molecule, which was postulated to be a similar salt formation reaction as observed for **31C**. However, the sample was not pure and could not be fully characterised, despite evidence for conversion of the sulfinamide starting material to product. Comparison of the sulfinamide starting material [**31L**] and the salt [**56L**] show significant differences in the aromatic region of the ^{1}H NMR (Figure 65).

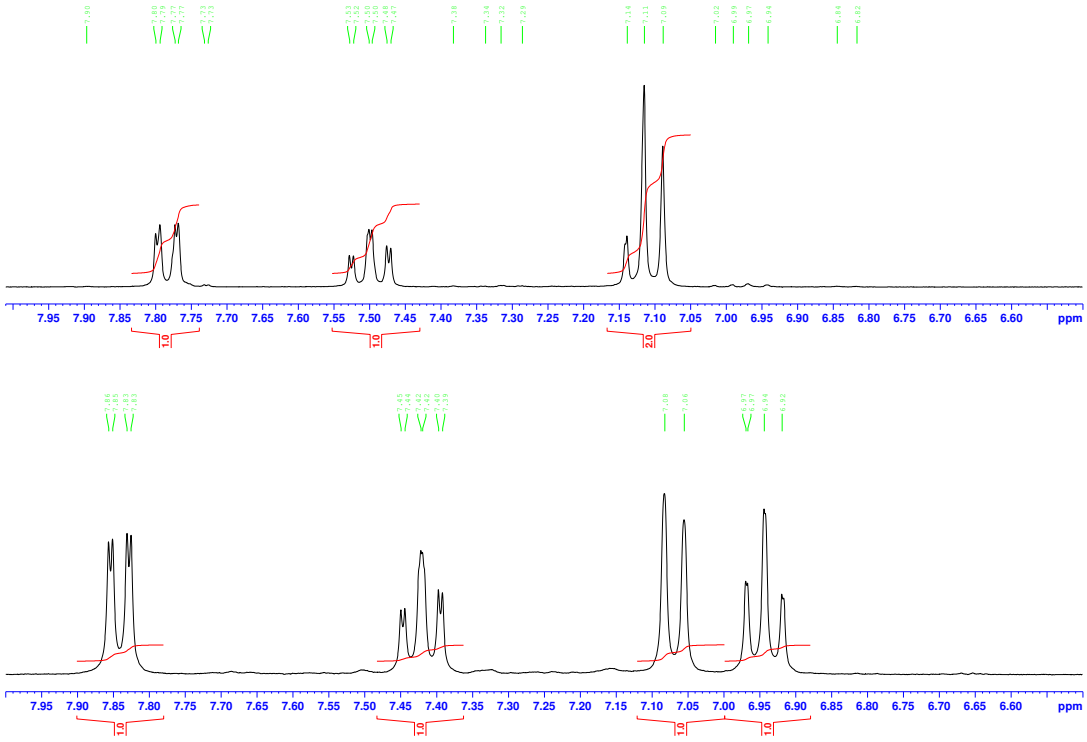


Figure 65: Comparison of the aromatic region in ^{1}H NMR spectra of 31L (top) and 56L [approx. 90% pure] (bottom).

The NMR spectrum of the sulfonate salt [**56L**] shows better peak resolution than that of the parent sulfinamide. All four aromatic proton signals are successfully resolved at 300 MHz due to the upfield shift of the ^{1}H multiplet that had previously overlapped with the doublet at approximately 7.07-7.17 ppm for the sulfinamide. The doublet of triplets is shifted upfield from 7.50 to 7.43 ppm and the doublet of doublets at 7.79 ppm is shifted downfield to 7.84 ppm. Efforts to fully characterise this material from ^{13}C NMR were unsuccessful due to strong signals from unidentified carbon sources.

After 3 weeks of stirring, efforts to prepare a pure sample of **56A**, **56M**, and **56O** were unsuccessful, with a complex, unidentifiable mixture recovered in each case.

2.5.4 Conclusions

The sulfinate and sulfonate salts display a very consistent pattern of N-H···O=S hydrogen bonds in a dependable and highly directional manner along the unit cell axes in all cases. The formation of hydrophilic and hydrophobic areas within the structures reflect that observed in the parent sulfinamides and also the analogous sulfonate salts observed in the literature. These interactions have contributed to the formation of good quality crystals that diffracted well in SCXRD studies, leading to easy determination of the crystal structures. Efforts to prepare and characterise pure samples of the salts were partially successful.

The study of sulfinamides and their crystal structures has presented significant, but not insurmountable challenges, while providing interesting results that have shed some valuable insight into the mechanistic pathways by which hydrolysis of these materials may occur. It seems that several factors are at play in determining the hydrolytic sensitivity of primary sulfinamides (mirrored by the debate in the literature), including the dihedral angle of the S=O to the aryl ring, the ability of the material to incorporate water into its crystal structure, and to a lesser extent, the electronic properties of the substituents on the ring.

The cocrystallization of sulfinamides was not pursued due to the observed instability of primary sulfinamides in solution. Efforts instead were directed toward the development of novel solid forms of other materials.

General Procedures

2.6 General Procedures 1: Organic Material Characterisation

2.6.1 Materials and Reagents

All reagents were of standard laboratory grade and were used without further purification. THF was distilled from sodium benzophenone ketyl before use,[101] all other solvents were of standard laboratory grade and used without further purification. *N*-Bromosuccinimide [**46**] was freshly recrystallised from hot water[25] and dried overnight before use; recrystallised material was stored in the freezer for use.

2.6.2 Chromatography

Flash column chromatography was performed using Kieselgel silica gel 60, 0.040-0.063 mm (Merck). Thin layer chromatography (TLC) was carried out on pre-coated silica gel plates (Merck 60 PF_{254}). Visualization was achieved by UV (254 nm) light detection in all cases.

2.6.3 Nuclear Magnetic Resonance (NMR)

Proton (300 MHz) NMR spectra were recorded on a Bruker Avance 300 MHz NMR spectrometer. Proton (600 MHz) and carbon (150 MHz) spectra were recorded on a Bruker Avance III 600 MHz NMR spectrometer using a 5mm Dual C-H cryoprobe. Spectra were recorded at room temperature (~20 ºC) in deuterated chloroform ($CDCl_3$) [internally referenced from trimethylsilane (TMS) chemical shift reference standard at 0.00 ppm], deuterated dimethylsulfoxide (DMSO-d_6) [internally referenced from the residual DMSO shift at 2.50 ppm], or deuterated methanol (CD_3OD) [internally referenced from residual methanol-shift at 3.31 ppm]. Chemical shifts (δ_H and δ_C) are reported in parts per million (ppm) relative to internal references [$CDCl_3$ at 77.16 ppm, DMSO-d_6 at 39.52 ppm, and CD_3OD at 49.00 ppm]. Coupling constants (*J*) are expressed in Hertz (Hz). Splitting patterns in ^1H spectra are designated as s (singlet), br s (broad singlet), d (doublet), t (triplet), q (quartet), dd (doublet of doublets), ddd (doublet of doublets of doublets), dt (doublet of triplets), and m (multiplet).

2.6.4 Infrared Spectroscopy (IR)

Infrared spectra were recorded on a Bruker Tensor 37 FT-IR spectrophotometer interfaced with Opus version 7.2.139.1294 over a range of 400 – 4000 cm^{-1}. An average of 16 scans was taken for each spectrum obtained with a resolution of 4 cm^{-1}. ν_{max} values were used where greater than 50% transmission was displayed [for bands of structural significance providing less than the above transmission the notation of weak (w) is applied]. Values are given to the nearest whole number in all cases.

2.6.5 Elemental Analysis (Microanalysis)

Elemental analyses were performed by the Microanalysis Laboratory, National University Ireland, Cork using either a Perkin-Elmer 240 or an Exeter Analytical CE440 elemental analyser.

2.6.6 Mass Spectrometry

Nominal mass spectra were recorded on a Waters Quattro Micro triple quadrupole spectrometer in electrospray ionisation (ESI) mode using 50% water/acetonitrile containing 0.1% formic acid as eluent; samples were made up in acetonitrile. High resolution mass spectra (HRMS) were recorded on a Waters LCT Premier Time of Flight spectrometer in electrospray ionisation (ESI) mode using 50% water/acetonitrile containing 0.1% formic acid as eluent; samples were made up in acetonitrile.

2.6.7 Treatment of Novel and Non-novel Materials

Novel materials for which no experimentally-determined spectral data was contained in the literature were fully characterised using proton and carbon NMR, infrared spectroscopy, mass spectrometry and either high-resolution mass spectrometry or elemental analysis. Non-novel materials were characterised using proton and carbon NMR, and infrared spectroscopy. Where spectral

characteristics for non-novel materials did not match with the literature data, materials were fully characterised.

2.7 General Procedures 2: Solid State Characterisation

2.7.1 Neat Grinding

Solid state grinding was performed using a Retsch MM400 Mixer Mill, equipped with two stainless steel 5 mL grinding vessels and one stainless steel grinding ball per vessel. All grinds were carried out at a rate of 30 Hz for 20 min using a 1:1 molar ratio of materials. Where an alternative ratio was determined from other analysis, neat grinding was repeated in 1:2 or 2:1 molar ratio, as appropriate.

2.7.2 Powder X-ray Diffraction (PXRD)

PXRD data was collected using either of the following instruments and parameters:

1. Bruker D2-Phaser bench-top powder X-ray diffractometer with Cu Kα radiation, using a LYNXEYE detector over the 2θ range of 3.5 – 45.5° with a step size of 0.075°. Samples were analysed neat in reflectance mode with no further sample preparation. Data collection parameters were set using the DIFFRAC.SUITE™ software and collected data was analysed using DIFFRAC.EVA™ software.
2. STOE STADI MP diffractometer with Cu Kα radiation using a linear PSD over the 2θ range of (3.5 – 50°), with a step size equal to 0.5 and step time equal to 30 second. The samples were prepared as transmission foils and the data was viewed via STOE WinXPOW POWDAT.

2.7.3 Differential Scanning Calorimetry (DSC)

DSC was collected using a TA instruments DSC Q1000 in conjunction with a refrigerated cooling system. The samples were equilibrated at 20 °C and ramped up to the required temperature at 5 °C per minute at a N_2 flow rate of 50 mL min^{-1}

(unless otherwise stated). The samples were prepared on aluminium pans and the data was viewed *via* TA Universal Analysis software.

2.7.4 Solution Crystallization

Solution crystallization experiments were conducted by dissolving the relevant material in approximately 10 mL of solvent in an 18 mm glass sample vial, after which the samples were allowed to stand, uncapped (unless otherwise specified) for complete evaporation of the solvent to occur.

2.7.5 Single Crystal X-ray Diffraction (SCXRD)

Single crystal X-ray data was collected on either a Bruker APEX II DUO diffractometer or a Bruker SMART X2S diffractometer[102] at temperatures between 100-300 K using graphite monochromatic Mo Kα (λ = 0.71073 Å) radiation. Calculations were performed using the APEX2 software suite.[103] The structures were solved using direct methods and refined on F^2 using SHELXL-97. Analysis was undertaken with the SHELX suite of programs and diagrams prepared with Mercury 3.5.1.[45,104] All non-hydrogen atoms were located and refined with anisotropic thermal parameters, unless otherwise specified. Hydrogen atoms were found and refined where possible; alternatively, hydrogens were included in calculated positions and allowed to ride on the parent atom for refinement.

2.7.6 Cambridge Structural Database[2] (CSD)

Searches of the Cambridge structural database were completed using Conquest version 1.18 as part of the CSD_2016 suite of programs. Searches of the CSD were performed with the requirement for organic material results only in all cases (organometallic results were excluded). In order to search for multi-component materials, the requirement for more than one unique molecule (Z >1) was included in the search term.

2.7.7 Density Functional Theory (DFT) Calculation

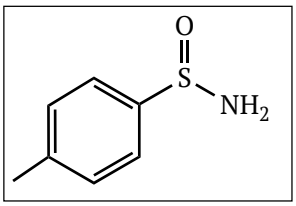
DFT calculations were performed using Spartan '14[105] to calculate molecular electrostatic potentials. Crystal Explorer 3.1[106] was used to generate 2-D plots and fingerprint diagrams.

Chapter 2

Experimental

2.8 Experimental

2.8.1 Attempted 'one-pot' Synthesis of 4-methylbenzene sulfinamide [31Q][11]



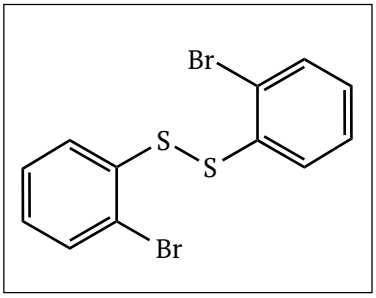
Attempted synthesis of 4-methylbenzene sulfinamide [**31Q**] was conducted according to the procedure outlined by Harmata *et al*.[11] 4-Methylbenzene sulfonyl chloride [**43**] (0.205 g, 1.0 mmol), was added to CH_2Cl_2 (3 mL) at 0 °C in a 50 mL round bottom flask equipped with a magnetic stir bar. With vigorous stirring, a solution of triethylamine (0.28 mL, 2.0 mmol), and triphenylphospine (0.262 g, 1.0 mmol) and aqueous ammonia solution (1.8 M, 0.55 mL, 1.0 mmol) in CH_2Cl_2 (3 mL) was added dropwise over a period of 1 h with vigorous stirring in between additions. After addition was complete, the reaction mixture was stirred for a further 30 min. The organic layer was washed with water (2x10 mL), dried (magnesium sulfate) and the solvent reduced *in vacuo*. The crude material showed no evidence of the desired product in NMR analysis.

2.8.2 Synthesis of Substituted Aryl Disulfides

General Procedure A:

Substituted aryl disulfides were prepared using a modified procedure from Kirihara *et al*.[16] Sodium iodide was replaced with potassium iodide in this work.

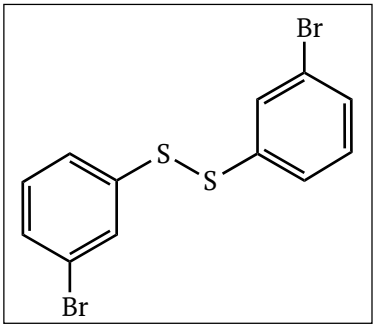
2-Bromobenzene disulfide [29C][17]



2-Bromobenzene thiol [**28C**] (3.347 g, 17.7 mmol) was dissolved in ethyl acetate (55 mL) in a 100 mL round bottom flask equipped with a magnetic stir bar. 30% Hydrogen peroxide [**37**] (1.95 mL, 17.7 mmol), and potassium iodide (0.029 g, 0.18 mmol) were added. The mixture was stirred at room

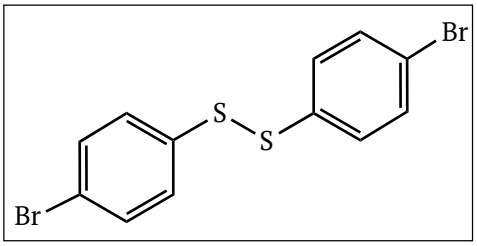
temperature for 30 min to yield an orange/brown solution. Addition of saturated sodium thiosulfate (40 mL) caused the reaction mixture to turn a pale yellow. Organic extractions with ethyl acetate (3 x 40 mL) were performed, and the combined organic portions were washed with brine (40 mL). The organic extracts were dried (magnesium sulfate), and the solvent was reduced *in vacuo*. The disulfide, **29C**, was isolated as a white solid (2.939 g, 88%) and was used without further purification; mp 96-99 °C (Lit.[17] 97 °C); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 7.10 (2H, td, *J* = 7.7, 1.5, 2xAr-**H**), 7.24-7.34 (2H, m, 2xAr-**H**), 7.56 (4H, dd, *J* = 8.1, 1.3, 4xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 121.3 (C, 2xAromatic-**C**Br), 127.2 (CH, 2xAromatic-**C**H), 128.1 (CH, 2xAromatic-**C**H), 128.4 (CH, 2xAromatic-**C**H), 133.1 (CH, 2xAromatic-**C**H), 136.3 (C, 2xAromatic-**C**S); m/z (ESI): 187 (^{79}Br), 189 (^{81}Br), ($C_6H_4BrS^-$) (1:0.75) [$(M/2)^-$]; ν_{max} (ATR)/cm^{-1}; 1440 (Aromatic C=C bend), 1036, 738, 704. Spectral characteristics are in agreement with those previously reported.[17]

3-Bromobenzene disulfide [29D]



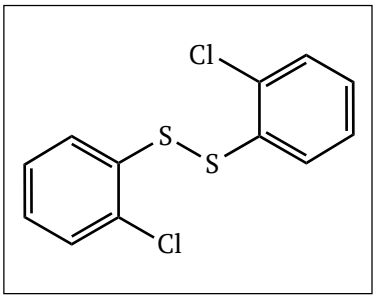
3-Bromobenzene thiol [**28D**] (5.159 g, 27.4 mmol), ethyl acetate (82 mL), 30% hydrogen peroxide [**37**] (3.00 mL, 27.3 mmol) and potassium iodide (0.045 g, 0.271 mmol) were used as described for **29C**. The disulfide, **29D**, was isolated as an orange oil (4.64 g, 91%), and was used without further purification; Microanalysis: $C_{12}H_8S_2Br_2$ requires C, 38.32; H, 2.14, S, 17.05%, Found: C 38.31, H 2.13, S, 16.85%; ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 7.20 (2H, t, *J* = 7.91, 2xAr-**H**), 7.34-7.46 (4H, m, 4xAr-**H**), 7.61-7.69 (2H, m, 2xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 123.3 (C, 2xAromatic-**C**Br), 126.1 (CH, 2xAromatic-**C**H), 130.1 (CH, 2xAromatic-**C**H), 130.62 (CH, 2xAromatic-**C**H), 130.64 (CH, 2xAromatic-**C**H), 138.8 (C, 2xAromatic-**C**S); m/z (ESI): 187 (^{79}Br), 189 (^{81}Br) (1:1) ($C_6H_4BrS^-$) [$(M/2)^-$]; ν_{max} (ATR)/cm^{-1}: 1570 (Aromatic C=C bend), 1559 (Aromatic C=C bend), 1454 (Aromatic C=C bend), 1080, 768, 741, 672, 651 (C-Br).

4-Bromobenzene disulfide [29E][18]



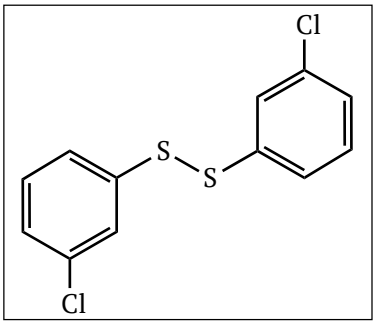
4-Bromobenzene thiol [**28E**] (4.203 g, 22.2 mmol), 30% hydrogen peroxide [**37**] (2.45 mL, 22.2 mmol), potassium iodide (0.037 g, 0.22 mmol), and ethyl acetate (67 mL) were used as described for **29C**. The disulfide, **29E**, was isolated as a yellow solid (3.523 g, 84%) and was used without further purification; mp 95-97 °C (Lit.[18] 97 °C); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 7.30-7.36 (4H, m, 4xAr-**H**), 7.40-7.46 (4H, m, 4xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 121.7 (C, 2xAromatic-**C**Br), 129.6 (CH, 4xAromatic-**C**H), 132.4 (CH, 4xAromatic-**C**H), 135.9 (C, 2xAromatic-**C**S); m/z (ESI): 187 (^{79}Br), 189 (^{81}Br) (1:1) ($C_6H_4BrS^-$) [$(M/2)^-$]; ν_{max} (ATR)/cm^{-1}: 1561 (w) (Aromatic C=C bend), 1004, 809, 625 (C-Br). Spectral characteristics are in agreement with those previously reported.[18]

2-Chlorobenzene disulfide [29F][17]



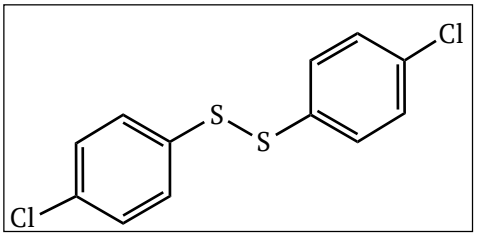
2-Chlorobenzene thiol [**28F**] (5.039 g, 34.8 mmol), ethyl acetate (105 mL), 30% hydrogen peroxide [**37**] (3.85 mL, 35.0 mmol) and potassium iodide (0.058 g, 0.35 mmol) were used as described for **29C**. The disulfide, **29F**, was isolated as an off-white solid (4.561 g, 91%), and was used without further purification; mp 83-85 °C (Lit.[17] 86-87 °C); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 7.11-7.26 (4H, m, 4xAr-**H**), 7.32-7.41 (2H, m, 2xAr-**H**), 7.51-7.59 (2H, m, 2xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 127.3 (CH, 2xAromatic-**C**H), 127.7 (CH, 2xAromatic-**C**H), 127.9 (CH, 2xAromatic-**C**H), 129.9 (CH, 2xAromatic-**C**H), 132.0 (C, 2xAromatic-**C**Cl), 134.5 (C, 2xAromatic-**C**S); m/z (ESI): 143 (^{35}Cl), 145 (^{37}Cl) (3:1) ($C_6H_4ClS^-$) [$(M/2)^-$]; ν_{max} (ATR)/cm^{-1}: 1446 (Aromatic C=C bend), 1430 (Aromatic C=C bend), 1028, 742 (C-Cl), 721, 660, 429. Spectral characteristics are in agreement with those previously reported.[17]

3-Chlorobenzene disulfide [29G][19]



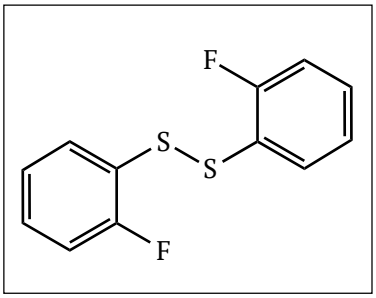
3-Chlorobenzene thiol [**28G**] (3.735 g, 26.0 mmol), ethyl acetate (77 mL), 30% hydrogen peroxide [**37**] (2.85 mL, 25.9 mmol) and potassium iodide (0.043 g, 0.26 mmol) were used as described for **29C**. The disulfide, **29G**, was isolated as an orange oil (3.671 g, 99%) [Lit.[19] colourless oil], and was used without further purification; ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 7.18-7.28 (4H, m, 4xAr**H**), 7.35 (2H, dt, *J* = 7.1, 1.7, 2xAr**H**), 7.46-7.49 (2H, m, 2xAr**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 125.5 (CH, 2xAromatic-**C**H), 127.1 (CH, 2xAromatic-**C**H), 127.7 (CH, 2xAromatic-**C**H), 130.4 (CH, 2xAromatic-**C**H), 135.3 (C, 2xAromatic-**C**Cl), 138.5 (C, 2xAromatic-**C**S); m/z (ESI): 143 (^{35}Cl), 145 (^{37}Cl) (3:1) ($C_6H_4ClS^-$) [$(M/2)^-$]; ν_{max} (ATR)/cm^{-1}; 1129, 957, 788 (C-Cl), 679, 659, 601. Spectral characteristics are in agreement with those previously reported.[19]

4-Chlorobenzene disulfide [29H][19]



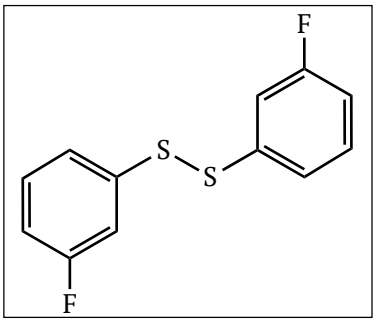
4-Chlorobenzene thiol [**28H**] (3.147 g, 21.8 mmol), ethyl acetate (65 mL), 30% hydrogen peroxide [**37**] (2.40 mL, 21.8 mmol) and potassium iodide (0.036 g, 0.22 mmol) were used as described for **29C**. The disulfide, **29H**, was isolated as a yellow solid (2.646 g, 85%) [Lit.[19] white solid], and was used without further purification; mp 73-75 °C (Lit.[19] 71-73 °C); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 7.27 (4H, d, *J* = 8.8, 4xAr-**H**), 7.40 (4H, d, *J* = 8.8, 4xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 129.4 (CH, 8xAromatic-**C**H), 133.8 (C, 2xAromatic-**C**Cl), 135.3 (C, 2xAromatic-**C**S); m/z (ESI): 143 (^{35}Cl), 145 (^{37}Cl) (3:1) ($C_6H_4ClS^-$) [$(M/2)^-$]; ν_{max} (ATR)/cm^{-1}; 1469 (w) (Aromatic C=C bend), 1112, 1008, 812, 740 (C-Cl). Spectral characteristics are in agreement with those previously reported.[19]

2-Fluorobenzene disulfide [29I][20]



2-Fluorobenzene thiol [**28I**] (4.235 g, 33.0 mmol), ethyl acetate (100 mL), 30% hydrogen peroxide [**37**] (3.6 mL, 30.0 mmol) and potassium iodide (0.055 g, 0.33 mmol) were used as described for **29C**he disulfide, **29I**, was isolated as a yellow oil (3.766 g, 90%) [Lit.[20] white solid], and was used without further purification; Microanalysis: $C_{12}H_8S_2F_2$ requires C, 56.68; H, 3.17, S, 25.21%, Found: C 56.88, H 3.03, S, 25.31%; ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 7.00-7.15 (4H, m, 4xAr-**H**), 7.20-7.32 (2H, m, 2xAr-**H**), 7.58 (2H, td, *J* = 7.7, 1.7, 2xAr-**H**), [Lit.[20] 7.07-7.13 (4H, m), 7.25-7.29 (2H, m), 7.60 (2H, t, *J* = 7.8)]; ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 115.9 (CH, d, *J* = 21.7, 2xAromatic-**C**H), 123.8 (C, d, *J* = 17.3, 2xAromatic-**C**S), 124.9 (CH, d, *J* = 3.8, 2xAromatic-**C**H), 129.9 (CH, d, *J* = 7.6, 2xAromatic-**C**H), 131.4 (CH, 2xAromatic-**C**H), 160.7 (C, d, *J* = 247.4, 2xAromatic-**C**F); m/z (ESI): 127 ($C_6H_4FS^-$) [$(M/2)^-$]; ν_{max} (ATR)/cm^{-1}: 1466 (Aromatic C=C bend), 1146 (Aromatic C=C bend), 1259 (C-F), 1220, 819, 746.

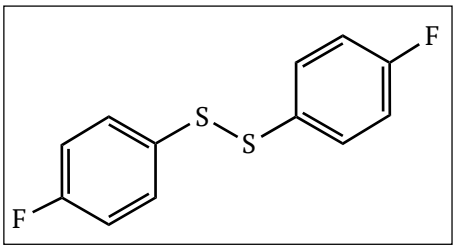
3-Fluorobenzene disulfide [29J]



3-Fluorobenzene thiol [**28J**] (3.823 g, 29.8 mmol), ethyl acetate (90 mL), 30% hydrogen peroxide [**37**] (3.30 mL, 30.0 mmol) and potassium iodide (0.049 g, 0.30 mmol) were used as described for **29C**. The disulfide, **29J**, was isolated as a yellow oil (3.285 g, 87%), and was used without further purification; ^{1}H NMR: δ_H (400 MHz) ($CDCl_3$): 6.85-6.98 (2H, m, 2xAr**H**), 7.16-7.32 (6H, m, 6xAr**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 114.2 (CH, d, *J* = 25.2, 2xAromatic-**C**H), 114.5 (CH, d, *J* = 22.3, 2xAromatic-**C**H), 112.8 (CH, d, *J* = 2.9, 2xAromatic-**C**H), 130.6 (CH, d, *J* = 8.4, 2xAromatic-**C**H), 138.9 (C, d, *J* = 7.2, 2xAromatic-**C**S), 163.2 (C, d, *J* = 249.1, 2xAromatic-**C**F); m/z (ESI): 253 [$(M-H)^-$]; HRMS (ESI): Exact mass calculated for $C_{12}H_9S_2F_2$ [$(M+H)^+$]: 255.0114, Found

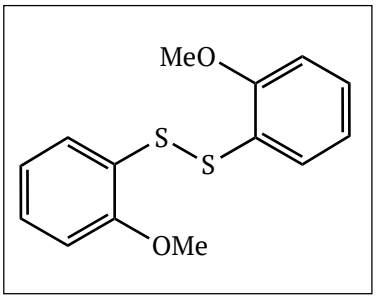
255.0102; ν_{max} (ATR)/cm^{-1}: 1578 (Aromatic C=C bend), 1469 (Aromatic C=C bend), 1429 (Aromatic C=C bend), 1298, 1263 (C-F), 872, 771, 673.

4-Fluorobenzene disulfide [29K][20]



4-Fluorobenzene thiol [**28K**] (4.102 g, 32.0 mmol), ethyl acetate (96 mL), 30% hydrogen peroxide [**37**] (3.50 mL, 31.8 mmol) and potassium iodide (0.053 g, 0.32 mmol), were used as described for **29C**he disulfide, **29K**, was isolated as a yellow oil (3.75 g, 92%) [Lit.[20] white solid], and was used without further purification; Microanalysis: $C_{12}H_8S_2F_2$ requires C, 56.68; H, 3.17, S, 25.21%, Found: C 56.71, H 3.37, S, 25.48%; ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 6.95-7.05 (4H, m, 4xAr**H**), 7.39-7.48 (4H, m, 4xAr**H**), [Lit.[20] 7.21 (4H, d, *J* = 8.7), 7.62 (4H, d, *J* = 8.7)]; ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 116.4 (CH, d, *J* = 22.4, 4xAromatic-**C**H), 131.4 (CH, d, *J* = 8.3, 4xAromatic-**C**H), 132.3 (C, d, *J* = 3.2, 2xAromatic-**C**S), 162.8 (C, d, *J* = 248.0, 2xAromatic-**C**F); m/z (ESI): 253 [$(M-H)^-$], 127 ($C_6H_4FS^-$) [$(M/2)^-$]; ν_{max} (ATR)/cm^{-1}: 1484 (Aromatic C=C bend), 1223 (C-F), 1153, 820, 620, 500 (S-S).

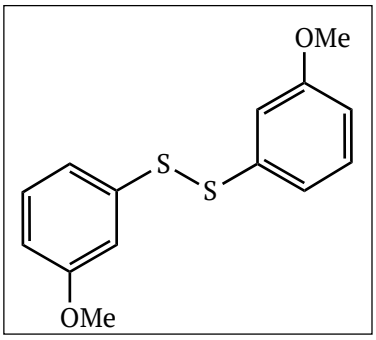
2-Methoxybenzene disulfide [29L][21]



2-Methoxybenzene thiol [**28L**] (5.132 g, 37.9 mmol), ethyl acetate (114 mL), 30% hydrogen peroxide [**37**] (4.20 mL, 38.2 mmol) and potassium iodide (0.063 g, 0.38 mmol), were used as described for **29C**. The disulfide, **29L**, was isolated as an off-white solid (4.79 g, 91%), and was used without further purification; mp 119-121 °C (Lit.[21] 118-119 °C); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.89 (6H, s, 2xO-C**H**$_3$), 6.78-6.97 (4H, m, 4xAr-**H**), 7.18 (2H, td, *J* = 7.8, 1.6, 2xAr-**H**), 7.53 (2H, dd, *J* = 7.8, 1.6, 2xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 56.0 (CH_3, 2xAr-O**C**H$_3$), 110.7 (CH, 2xAromatic-**C**H), 121.5 (CH, 2xAromatic-**C**H), 124.8 (C, 2xAromatic-**C**S), 127.8 (CH, 2xAromatic-**C**H), 127.9

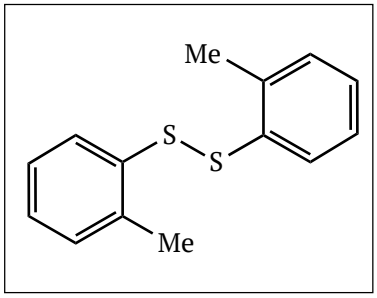
(CH, 2xAromatic-**C**H), 156.8 (C, 2xAromatic-**C**OCH_3); m/z (ESI): 301 [$(M+Na)^+$]; ν_{max} (ATR)/cm^{-1}: 1472 (Aromatic C=C bend), 1463 (Aromatic C=C bend), 1270, 1235 (C-O), 1058, 1037, 739. Spectral characteristics are in agreement with those previously reported.[21]

3-Methoxybenzene disulfide [29M][19,20]



3-Methoxybenzene thiol [**28M**] (5.108 g, 36.4 mmol) in ethyl acetate (110 mL), 30% hydrogen peroxide [**37**] (4.0 mL, 36.4 mmol) and potassium iodide (0.060 g, 0.36 mmol) were used as described for **29C**. The disulfide, **29M**, was isolated as a red-brown oil (4.872 g, 96%) [Lit.[20] white solid], and was used without further purification; ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.75 (6H, s, 2xAr-C**H**$_3$), 6.70-6.79 (2H, m, 2xAr-**H**) 7.02-7.23 (4H, m, 4xAr-**H**), 7.24-7.27 (2H, m, 2xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 55.4 (CH_3, 2xAr-O**C**H_3), 112.8 (CH, 2xAromatic-**C**H), 113.3 (CH, 2xAromatic-**C**H), 119.8 (CH, 2xAromatic-**C**H), 130.0 (CH, 2xAromatic-**C**H), 138.4 (C, 2xAromatic-**C**S), 160.2 (C, 2xAromatic-**C**OCH_3); m/z (ESI): 279 [$(M+H)^+$], 301 [$(M+Na)^+$], 139 ($C_7H_7S^-$) [$(M/2)^-$]; ν_{max} (ATR)/cm^{-1}: 1587 (Aromatic C=C bend), 1573 (Aromatic C=C bend), 1474 (Aromatic C=C bend), 1422 (Aromatic C=C bend), 1281, 1244 (C-O), 1227, 1070, 854, 841, 767, 682. Spectral characteristics are in agreement with those previously reported.[19]

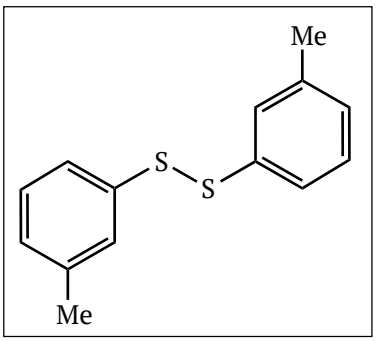
2-Methylbenzene disulfide [29O][22,107]



2-Methylbenzene thiol [**28O**] (2.910 g, 23.4 mmol), ethyl acetate (70 mL), 30% hydrogen peroxide [**37**] (2.60 mL, 23.4 mmol) and potassium iodide (0.039 g, 0.23 mmol) were used as described for **29C**. The disulfide, **29O**, was isolated as a yellow solid (2.57 g, 89%), and was used without further purification;

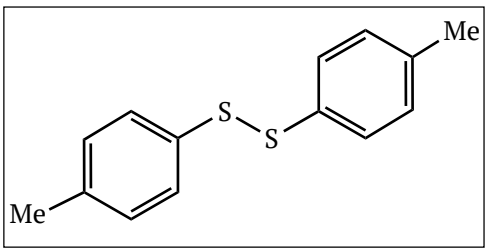
mp 36-39 °C (Lit.[107] 36-38 °C); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$); 2.42 (6H, s, 2xAr-C**H**$_3$), 7.07-7.20 (6H, m, 6xAr-**H**), 7.47-7.55 (2H, m, 2xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 20.1 (CH_3, 2xAromatic-**C**H$_3$), 126.8 (CH, 2xAromatic-**C**H), 127.5 (CH, 2xAromatic-**C**H), 128.8 (CH, 2xAromatic-**C**H), 130.4 (CH, 2xAromatic-**C**H), 135.6 (C, 2xAromatic-**C**CH$_3$), 137.5 (C, 2xAromatic-**C**S); m/z (ESI): 245 [(M-H)$^-$], 123 ($C_7H_7S^-$) [(M/2)$^-$]; ν_{max} (ATR)/cm^{-1}: 1462 (Aromatic C=C bend), 1054, 734, 674. Spectral characteristics are in agreement with those previously reported.[22]

3-Methylbenzene disulfide [29P][23]



3-Methylbenzene thiol [**28P**] (2.302 g, 18.5 mmol), ethyl acetate (56 mL), 30% hydrogen peroxide [**37**] (2.05 mL, 18.6 mmol) and potassium iodide (0.031 g, 0.19 mmol) were used as described for **29C**. The disulfide, **29P**, was isolated as a yellow oil (2.045 g, 90%), and was used without further purification; ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 2.30 (6H, s, 2xAr-C**H**$_3$), 7.01 (2H, d, *J* = 7.8, 2xAr-**H**), 7.12-7.23 (2H, m, 2xAr-**H**), 7.25-7.35 (4H, m, 4xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 21.5 (CH_3, 2xAromatic-**C**H$_3$), 124.7 (CH, 2xAromatic-**C**H), 128.1 (CH, 2xAromatic-**C**H), 128.2 (CH, 2xAromatic-**C**H), 129.0 (CH, 2xAromatic-**C**H), 137.1 (C, 2xAromatic-**C**CH$_3$), 139.0 (C, 2xAromatic-**C**S); m/z (ESI): 245 [(M-H)$^-$], 123 ($C_7H_7S^-$) [(M/2)$^-$]; ν_{max} (ATR)/cm^{-1}: 1591 (Aromatic C=C bend), 1471 (Aromatic C=C bend), 769, 684. Spectral characteristics are in agreement with those previously reported.[23]

4-Methylbenzene disulfide [29Q][16]



4-Methylbenzene thiol [**28Q**] (1.512 g, 12.2 mmol), ethyl acetate (37 mL), 30% hydrogen peroxide [**37**] (1.34 mL, 12.2 mmol) and potassium iodide (0.020 g, 0.12 mmol) were used as described for

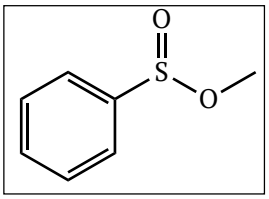
29C. The disulfide, **29Q**, was isolated as a yellow solid (1.079 g, 72%), and was used without further purification; mp 44-46 °C (Lit.[16] 44-45 °C); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 2.32 (6H, s, 2xC**H**$_3$), 7.11 (4H, d, *J* = 8.0, 4xAr-**H**), 7.39 (4H, d, *J* = 8.2, 4xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 21.2 (CH_3, 2xAromatic-**C**H_3), 128.7 (CH, 4xAromatic-**C**H), 129.9 (CH, 4xAromatic-**C**H), 134.0 (C, 2xAromatic-**C**CH_3), 137.5 (C, 2xAromatic-**C**S); m/z (ESI): 123 ($C_7H_7S^-$) [$(M/2)^-$]; ν_{max} (ATR)/cm^{-1}: 1487 (Aromatic C=C bend), 1014 (w), 801. Spectral characteristics are in agreement with those previously reported.[16]

2.8.3 Synthesis of Substituted Methyl Aryl Sulfinate Esters

General Procedure B:

Substituted aromatic methyl sulfinate esters were prepared using the synthetic methodology outlined by Brownbridge and Jowett.[15]

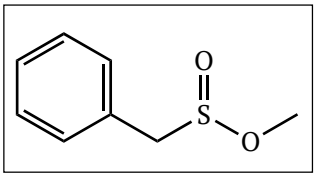
Methyl benzene sulfinate [30A][26]



Methyl benzene sulfinate [**30A**] was prepared by dissolving diphenyl disulfide (**29A**) (1.792 g, 8.2 mmol) in a mixture of methanol (41 mL) and CH_2Cl_2 (41 mL) in a 100 mL round bottom flask equipped with a magnetic stir bar. The flask was cooled to 0 °C in an ice-bath. *N*-Bromosuccinimide [**46**] (4.382 g, 24.6 mmol) was added to the stirring solution, at which time a bright orange colour gradually became apparent (over 2-5 min). The reaction was allowed to warm slowly to room temperature and reaction progress was monitored via TLC for the disappearance of the disulfide band. The reaction mixture was transferred to a separatory funnel and a further portion of CH_2Cl_2 (41 mL) was added. The organic mixture was washed with saturated sodium bicarbonate solution (3 x 45 mL), and water (45 mL), and dried (magnesium sulfate). The solvent was removed under reduced pressure. The sulfinate ester, **30A**, was isolated as a colourless oil (1.567 g, 61%), and was used without further purification; ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.49 (3H, s, S-OC**H**$_3$), 7.51-7.61 (3H, m, 3xAr-**H**), 7.68-7.75 (2H, m, 2xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 49.8

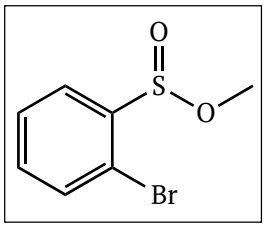
(CH_3, S-O**C**H_3), 125.5 (CH, 2xAromatic-**C**H), 129.2 (CH, 2xAromatic-**C**H), 132.4 (CH, Aromatic-**C**H), 144.0 (C, Aromatic-**C**S); ν_{max} (ATR)/cm^{-1}: 1124 (S=O), 958 (S-O), 755, 689, 673. Spectral characteristics are in agreement with those previously reported.[26]

Methyl phenylmethane sulfinate [30B][15]



Dibenzyl disulfide [**29B**] (1.462 g, 5.93 mmol), methanol (30 mL), CH_2Cl_2 (30 mL), and *N*-bromosuccinimide [**46**] (3.168 g, 17.8 mmol) were used as described for **30A**. The sulfinate ester, **30B**, was isolated as a colourless oil (1.7135 g, 85%) and was used without further purification; ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.74 (3H, s, C**H**$_3$), 4.00 (2H, ABq, *J* = 15.3, Ar-C**H**$_2$), 7.23-7.43 (5H, m, 5xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 54.8 (CH_3, S-O**C**H_3), 64.1 (CH_2, Aromatic-**C**H_2), 128.4 (CH, Aromatic-**C**H), 128.8 (C, Aromatic-**C**S), 128.9 (CH, 2xAromatic-**C**H), 130.5 (CH, 2xAromatic-**C**H); m/z (ESI): 171 [$(M+H)^+$]; HRMS (ESI): Exact mass calculated for $C_8H_{11}O_2S$ [$(M+H)^+$] 171.0480, Found 171.0473; ν_{max} (ATR)/cm^{-1}: 1112 (S=O), 987 (S-O), 770, 680, 632. Spectral characteristics are in agreement with those previously reported.[15]

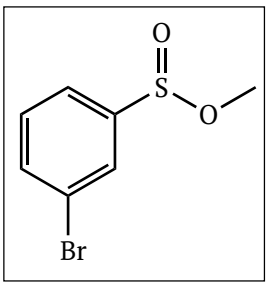
Methyl 2-bromobenzene sulfinate [30C]



2-Bromobenzene disulfide [**29C**] (2.755 g, 7.3 mmol), methanol (35 mL), CH_2Cl_2 (35 mL) and *N*-bromosuccinimide [**46**] (3.911 g, 21.9 mmol) were used as described for **30A**. Column chromatography on silica gel (90:10 hexane/ethyl acetate) afforded the pure sulfinate ester, **30C**, as a yellow oil (2.334 g, 68%); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.60 (3H, s, S-OC**H**$_3$), 7.42 (1H, td, *J* = 7.6, 1.8, Ar-**H**), 7.54 (1H, td, *J* = 7.6, 1.2, Ar-**H**), 7.63 (1H, dd, *J* = 7.8, 1.2, Ar-**H**), 7.94 (1H, dd, *J* = 7.7, 1.7, Ar-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 56.8 (CH_3, S-O**C**H_3), 121.0 (C, Aromatic-**C**Br), 127.8, 132.4, 134.9 (CH, 3xAromatic-**C**H), 135.4 (C, Aromatic-**C**S), 135.9 (CH, Aromatic-**C**H); m/z

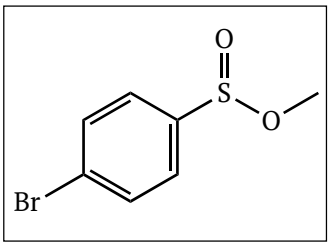
(ESI): 235 (^{79}Br), 237 (^{81}Br) (1:1) [(M+H)$^+$], 257 (^{79}Br), 259 (^{81}Br) (1:1) [(M+Na)$^+$]; HRMS (ESI): Exact mass calculated for $C_7H_8O_2SBr$ (^{79}Br) [(M+H)$^+$]: 234.9428, Found 234.9437, $C_7H_8O_2SBr$ (^{81}Br) [(M+H)$^+$]: 236.9408, Found 236.9415; ν_{max} (ATR)/cm^{-1}; 1359, 1281, 1183 (S=O), 977 (S-O), 757, 725, 705.

Methyl 3-bromobenzene sulfinate [30D]



3-Bromobenzene disulfide [**29D**] (3.872 g, 10.3 mmol), methanol (50 mL), CH_2Cl_2 (50 mL) and *N*-bromosuccinimide [**46**] (5.493 g, 30.9 mmol), were used as described for **30A** in a 250 mL round bottom flask. The sulfinate ester, **30D**, was isolated as an orange oil (3.423 g, 71%), and was used without further purification; ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.51 (3H, s, S-OC**H**$_3$), 7.43 (1H, t, *J* = 7.8, Ar-**H**), 7.60-7.66 (1H, m, Ar-**H**), 7.66-7.72 (1H, m, Ar-**H**), 7.84-7.87 (1H, m, Ar-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 50.1 (CH_3, S-O**C**H$_3$), 123.5 (C, Aromatic-**C**Br), 124.2, 128.5, 130.8, 135.4 (CH, 4xAromatic-**C**H), 146.1 (C, Aromatic-**C**S); m/z (ESI): 235 (^{79}Br), 237 (^{81}Br) (1:1) [(M+H)$^+$]; HRMS (ESI): Exact mass calculated for $C_7H_8O_2SBr$ (^{79}Br) [(M+H)$^+$] 234.9428, Found 234.9422; ν_{max} (ATR)/cm^{-1}: 1128 (S=O), 956 (S-O), 786, 678, 649. ^{1}H NMR spectrum is consistent with that in the literature.[35]

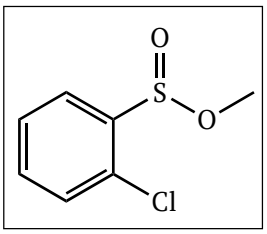
Methyl 4-bromobenzene sulfinate [30E][27]



4-Bromobenzene disulfide [**29E**] (3.014 g, 8.0 mmol), methanol (40 mL), CH_2Cl_2 (40 mL), and *N*-bromosuccinimide [**46**] (4.279 g, 24.0 mmol) were used as described for **30A**. The pure sulfinate ester, **30E**, was isolated as a colourless oil (2.731 g, 72.5%), and was used without further purification; ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.49 (3H, s, S-OC**H**$_3$), 7.54-7.61 (2H, m, 2xAr-**H**), 7.66-7.73 (2H, m, 2xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 50.0 (CH_3, S-O**C**H$_3$), 126.7 (C, Aromatic-**C**Br), 127.2 (CH, 2xAromatic-**C**H), 132.5 (CH, 2xAromatic-**C**H), 143.2 (C, Aromatic-**C**S); m/z (ESI):

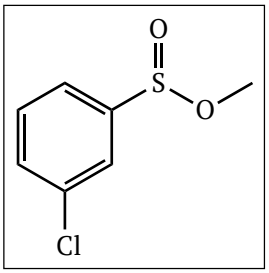
235 (^{79}Br), 237 (^{81}Br) (1:1) [(M+H)$^+$]; HRMS (ESI); Exact mass calculated for $C_7H_8O_2SBr$ (^{79}Br) [(M+H)$^+$] 234.9428, Found 234.9428; ν_{max} (ATR)/cm^{-1}; 1129 (S=O), 1062, 1008, 954 (S-O), 818, 726, 705, 679. At time of preparation, no characteristic data was contained in the literature for this material, therefore full characterisation was performed. Spectral characteristics are in agreement with those published in 2016.[27]

Methyl 2-chlorobenzene sulfinate [30F][28]



2-Chlorobenzene disulfide [**29F**] (2.359 g, 8.3 mmol), methanol (41 mL), CH_2Cl_2 (42 mL) and *N*-bromosuccinimide [**46**] (4.431 g, 24.9 mmol) were used as described for **30A**. Column chromatography on silica gel (CH_2Cl_2 eluent) afforded the sulfinate ester, **30F**, as a colourless oil (2.561 g, 81%); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.59 (3H, s, S-OC**H**$_3$), 7.41-7.55 (3H, m, 3xAr-**H**), 7.90-7.98 (1H, m, Ar-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 51.4 (C, S-O**C**H$_3$), 126.8, 127.3, 130.5 (CH, 3xAromatic-**C**H), 132.8 (C, Aromatic-**C**Cl), 133.6 (CH, Aromatic-**C**H), 141.1 (C, Aromatic-**C**S); ν_{max} (ATR)/cm^{-1}: 1124 (S=O), 1034, 962 (S-O), 760, 733, 714, 679; Spectral characteristics are in agreement with those previously reported.[28]

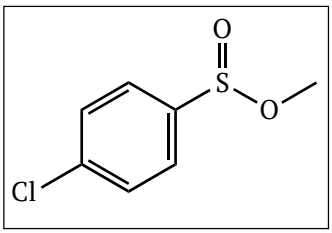
Methyl 3-chlorobenzene sulfinate [30G][27]



3-Chlorobenzene disulfide [**29G**] (2.978 g, 10.4 mmol), methanol (51 mL), CH_2Cl_2 (50 mL), and *N*-bromosuccinimide [**46**] (5.561 g, 31.2 mmol) were used as described for **30A**. Column chromatography on silica gel (CH_2Cl_2 eluent) afforded the sulfinate ester, **30G**, as a colourless oil (2.230 g, 56%); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.51 (3H, s, S-OC**H**$_3$), 7.45-7.62 (3H, m, 3x Ar-**H**), 7.67-7.73 (1H, m, Ar-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 50.0 (C, S-O**C**H$_3$), 123.7, 125.7, 130.5, 132.5 (CH, 4xAromatic-**C**H), 135.6 (C, Aromatic-**C**Cl), 146.0 (C, Aromatic-**C**S); m/z

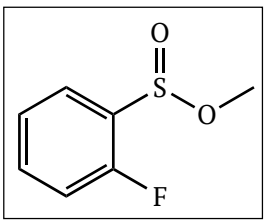
(ESI): 191 (^{35}Cl), 193 (^{37}Cl) (3:1) [(M+H)$^+$]; HRMS (ESI): Exact mass calculated for $C_7H_8O_2SCl$ (^{35}Cl) [(M+H)$^+$] 190.9934, Found 190.9931; ν_{max} (ATR)/cm^{-1}: 1129 (S=O), 957 (S-O), 789, 680, 659. At time of preparation, no characteristic data was contained in the literature for this material, therefore full characterisation was performed. Spectral characteristics are in agreement with those published in 2016.[27]

Methyl 4-chlorobenzene sulfinate [30H][29]



4-Chlorobenzene disulphide [**29H**] (2.509 g, 8.8 mmol), methanol (44 mL), CH_2Cl_2 (44 mL) and *N*-bromosuccinimide [**46**] (4.682 g, 26.3 mmol) were used as described for **30A**. Column chromatography on silica gel (90:10 hexane/ethyl acetate eluent) afforded the sulfinate ester, **30H**, as a yellow oil (2.982 g, 89%); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.49 (3H, s, S-OC**H**$_3$), 7.49-7.59 (2H, m, 2xAr-**H**), 7.61-7.69 (2H, m, 2xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 49.9 (C, S-O**C**H$_3$), 127.1 (CH, 2xAromatic-**C**H), 129.8 (CH, 2xAromatic-**C**H), 138.8 (C, Aromatic-**C**Cl), 142.6 (C, Aromatic-**C**S); m/z (ESI): 191 (^{35}Cl), 193 (^{37}Cl) (3:1) [(M+H)$^+$]; HRMS (ESI): Exact mass calculated for $C_7H_8O_2SCl$ (^{35}Cl) [(M+H)$^+$] 190.9934, Found 190.9935; ν_{max} (ATR)/cm^{-1}: 1125 (w) (S=O), [Lit.[29] 1135], 1085, 1032, 958 (S-O), 822, 708. Spectral characteristics are consistent with those previously reported for (R_s)-methyl 4-chlorobenzene sulfinate.[29]

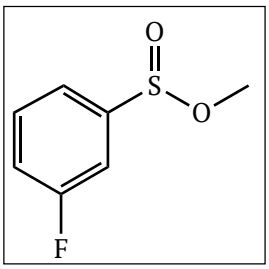
Methyl 2-fluorobenzene sulfinate [30I]



2-Fluorobenzene disulfide [**29I**] (3.266 g, 12.8 mmol) in methanol (60 mL), CH_2Cl_2 (60 mL), and *N*-bromosuccinimide [**46**] (6.849 g, 38.5 mmol), were used as described for **30A**. Column chromatography on silica gel (CH_2Cl_2 eluent) afforded the sulfinate ester, **30I**, as a colourless oil (2.445 g, 59%); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.59 (3H, s,

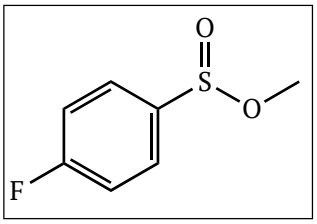
S-OC**H**$_3$) 7.12-7.21 (1H, m, Ar-**H**), 7.36 (1H, td, *J* = 7.6, 1.0, Ar-**H**), 7.51-7.62 (1H, m, Ar-**H**) 7.81-7.90 (1H, m, Ar-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 51.0 (CH_3, S-O**C**H$_3$), 116.4 (CH, d, *J* = 19.7, Aromatic-**C**H), 124.7 (CH, d, *J* = 3.6, Aromatic-**C**H), 126.9 (CH, d, *J* = 2.1, Aromatic-**C**H), 131.4 (C, d, *J* = 15.9, Aromatic-**C**S), 134.5 (CH, d, *J* = 7.8, Aromatic-**C**H), 160.0 (C, d, *J* = 252.0, Aromatic-**C**F); m/z (ESI): 175 [(M+H)$^+$], 197 [(M+Na)$^+$]; HRMS (ESI): Exact mass calculated for $C_7H_8O_2SF$ [(M+H)$^+$] 175.0229, Found 175.0226; ν_{max} (ATR)/cm^{-1}: 1469 (Aromatic C=C bend), 1133 (S=O), 1120, 958 (S-O), 820, 722, 682.

Methyl 3-fluorobenzene sulfinate [30J]



3-Fluorobenzene disulfide [**29J**] (0.942 g, 3.7 mmol), methanol (37 mL), CH_2Cl_2 (37 mL) and *N*-bromosuccinimide [**46**] (1.98 g, 11.1 mmol) were used as described for **30A**. The sulfinate ester, **30J**, was isolated as a colourless oil (1.145 g, 89%) and was used without further purification; ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.50 (3H, s, S-OC**H**$_3$) 7.20-7.32 (1H, m, Ar-**H**), 7.39-7.60 (3H, m, 3xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$); 50.0 (CH_3, S-O**C**H$_3$), 112.8 (CH, d, *J* = 23.4, Aromatic-**C**H), 119.5 (CH, d, *J* = 21.5, Aromatic-**C**H), 121.4 (CH, d, *J* = 3.3, Aromatic-**C**H), 131.0 (CH, d, *J* = 7.8, Aromatic-**C**H), 146.6 (C, d, *J* = 6.3, Aromatic-**C**S), 163.1 (C, d, *J* = 252.5, Aromatic-**C**F); m/z (ESI): 175 [(M+H)$^+$], 197 [(M+Na)$^+$]; HRMS (ESI): Exact mass calculated for $C_7H_8O_2SF$ [(M+H)$^+$] 175.0229, Found 175.0230; ν_{max} (ATR)/cm^{-1}: 1471 (Aromatic C=C bend), 1216 (Aromatic C-F), 1127 (S=O), 958 (S-O), 874, 789, 681, 613.

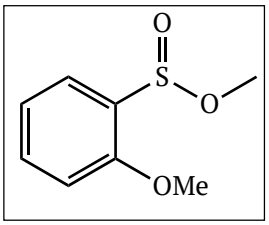
Methyl 4-fluorobenzene sulfinate [30K][26]



4-Fluorobenzene disulfide [**29K**] (3.26 g, 12.8 mmol), methanol (65 mL), CH_2Cl_2 (65 mL), and *N*-bromosuccinimide [**46**] (6.84 g, 38.4 mmol) were used as described for **30A**. Column chromatography on

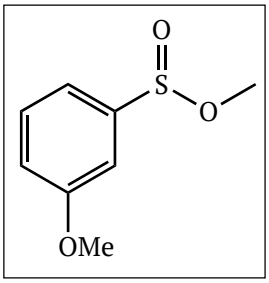
silica gel (90:10 hexane/ethyl acetate eluent) afforded the sulfinate ester, **30K**, as a yellow oil (2.22 g, 50%); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.49 (3H, s, C**H**$_3$), 7.19-7.29 (2H, m, 2xAr-**H**), 7.67-7.76 (2H, m, 2xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 49.8 (CH_3, S-O**C**H_3), 116.5 (CH, d, *J* = 22.7, 2xAromatic-**C**H), 128.0 (CH, d, *J* = 9.0, 2xAromatic-**C**H), 140.0 (C, d, *J* = 3.0, Aromatic-**C**S), 165.2 (C, d, *J* = 253.1, Aromatic-**C**F); ν_{max} (ATR)/cm^{-1}: 1588 (Aromatic C=C bend), 1490 (Aromatic C=C bend), 1227 (Aromatic C-F), 1127 (S=O), 1076, 1011, 958 (S-O), 834, 815, 675, 659, 634. Spectral characteristics are in agreement with those previously reported.[26]

Methyl 2-methoxybenzene sulfinate [30L][14]



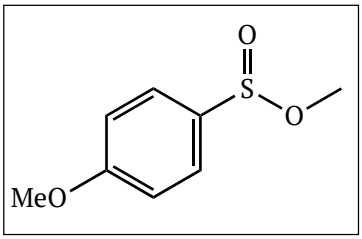
2-Methoxybenzene disulfide [**29L**] (3.727 g, 13.4 mmol), methanol (67 mL), CH_2Cl_2 (67 mL), and *N*-bromosuccinimide [**46**] (7.151 g, 40.2 mmol) were used as described for **30A**. Column chromatography on silica gel (80:20 hexane:ethyl acetate eluent) afforded the sulfinate ester, **30L**, as a pale yellow oil (3.859 g, 77%); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.53 (3H, s, S-OC**H**$_3$), 3.91 (3H, s, Ar-OC**H**$_3$), 6.97 (1H, d, *J* = 8.2, Ar-**H**), 7.13 (1H, td, *J* = 7.5, 0.9, Ar-**H**), 7.47-7.57 (1H, m, Ar-**H**), 7.84 (1H, dd, *J* = 7.7, 1.7, Ar-**H**); [Lit.[14] 3.36 (3H, s, C**H**$_3$), 3.75 (3H, s, C**H**$_3$), 6.85 (1H, dd, *J* = 7.4, 0.8), 6.98 (1H, dt, *J* = 6.7, 0.9), 7.42-7.33 (1H, m), 7.66 (dd, J = 5.9, 1.7 Hz, 1H)]; ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 50.5 (CH_3, S-O**C**H_3), 56.1 (CH_3, Ar-O**C**H_3), 111.5, 120.9, 126.5, 131.3 (CH, 4xAromatic-**C**H), 134.1 (C, Aromatic-**C**S), 157.7 (C, Aromatic-**C**OCH_3), [Lit.[14] 49.8, 55.5, 111.1, 120.2, 125.5 (2xC), 133.6, 157.1]; m/z (ESI): 187 [$(M+H)^+$], 209 [$(M+Na)^+$]; HRMS (ESI): Exact mass calculated for $C_8H_{11}O_3S$ [$(M+H)^+$] 187.0429, Found 187.0422; ν_{max} (ATR)/cm^{-1}; 1476 (Aromatic C=C bend), 1273 (C-O), 1124 (S=O), 1118 (C-O), 1017, 962 (S-O), 795, 756, 671; [Lit.[14] 3009, 2975, 1737, 1436]. The spectral characteristics recorded are in agreement with those previously reported, although a slight shift is noted for the NMR spectra.[14]

Methyl 3-methoxybenzene sulfinate [30M]



3-Methoxybenzene disulfide [**29M**] (3.170 g, 11.4 mmol), methanol (67 mL), CH_2Cl_2 (67 mL), and *N*-bromosuccinimide [**46**] (6.082 g, 34.2 mmol) were used as described for **30A**. Column chromatography on silica gel (80:20 hexane:ethyl acetate eluent) afforded the sulfinate ester, **30M**, as a pale yellow oil (3.800 g, 90%); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.49 (3H, s, S-OC**H**$_3$), 3.87 (3H, s, Ar-OC**H**$_3$), 7.08 (1H, ddd, *J* = 8.2, 2.6, 1.0, Ar-**H**), 7.22-7.29 (2H, m, 2xAr-**H**), 7.40-7.49 (1H, m, Ar-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 49.8 (C, S-O**C**H$_3$), 55.7 (CH$_3$, Ar-O**C**H$_3$), 109.8, 117.8, 118.9, 130.3 (CH, 4xAromatic-**C**H), 145.5 (C, Aromatic-**C**S), 160.4 (CH, Aromatic-**C**OCH$_3$); m/z (ESI): 187 [(M+H)$^+$], 209 [(M+Na)$^+$]; HRMS (ESI): Exact mass calculated for $C_8H_{11}O_3S$ [(M+H)$^+$] 187.0429, Found 187.0429; ν_{max} (ATR)/cm^{-1}; 1476 (Aromatic C=C bend), 1246 (C-O) 1235 (C-O), 1125 (S=O), 1034, 958 (S-O), 785, 677.

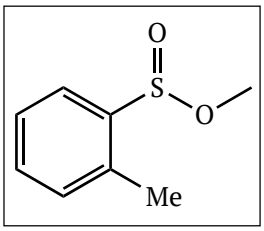
Methyl 4-methoxybenzene sulfinate [30N][27]



4-Methoxybenzene disulfide [**29N**] (2.583 g, 9.3 mmol), methanol (47 mL), CH_2Cl_2 (46 mL), and *N*-bromosuccinimide [**46**] (4.962 g, 27.9 mmol) were used as described for **30A**. Column chromatography on silica gel (90:10 CH_2Cl_2:hexane eluent) afforded the sulfinate ester, **30N**, as a pale yellow oil (1.282 g, 74%); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.46 (3H, s, S-OC**H**$_3$), 3.87 (3H, s, Ar-OC**H**$_3$), 7.00-7.07 (2H, m, 2xAr-**H**), 7.60-7.68 (2H, m, 2xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 49.4 (C, S-O**C**H$_3$), 55.7 (CH$_3$, Ar-O**C**H$_3$), 114.6 (CH, 2xAromatic-**C**H), 127.4 (CH, 2xAromatic-**C**H), 135.7 (C, Aromatic-**C**S), 162.9 (CH, Aromatic-**C**OCH$_3$); m/z (ESI): 187 [(M+H$^+$)], 209 [(M+Na)$^+$]; HRMS (ESI): Exact mass calculated for $C_8H_{11}O_3S$ [(M+H)$^+$] 187.0429, Found 187.0425; ν_{max} (ATR)/cm^{-1}; 1577 (Aromatic C=C bend), 1494 (Aromatic C=C bend), 1252 (C-O), 1123 (S=O), 1104, 1024, 957 (S-O), 831, 797, 671. At time of preparation, no characteristic data was contained

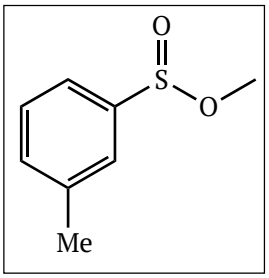
in the literature for this material, therefore full characterisation was performed. Spectral characteristics are in agreement with those published in 2016.[27]

Methyl 2-methylbenzene sulfinate [30O][28]



2-Methylbenzene disulfide [**29O**] (1.792 g, 7.3 mmol), methanol (37 mL), CH_2Cl_2 (36 mL), and *N*-bromosuccinimide [**46**] (3.879 g, 21.8 mmol) were used as described for **30A**. Column chromatography on silica gel (CH_2Cl_2eluent) afforded the sulfinate ester, **30O**, as a pale yellow oil (1.768 g, 71%); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 2.49 (3H, s, Ar-C**H**$_3$), 3.48 (3H, s, S-OC**H**$_3$), 7.21-7.29 (1H, m, Ar-**H**), 7.35-7.46 (2H, m, 2xAr-**H**), 7.90 (1H, dd, *J* = 7.3, 1.8, Ar-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 18.1 (CH_3, Aromatic-**C**H$_3$), 50.0 (CH_3, S-O**C**H$_3$), 124.8, 126.4, 131.3, 132.3 (CH, 4xAromatic-**C**H), 136.8 (C, Aromatic-**C**CH$_3$), 141.4 (C, Aromatic-**C**S); ν_{max} (ATR)/cm^{-1}: 1123 (S=O), 962 (S-O), 714, 681, 668; Spectral characteristics are in agreement with those previously reported.[28]

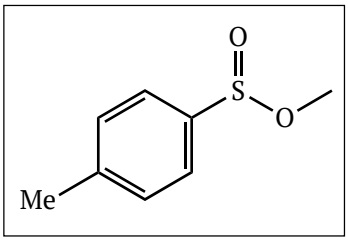
Methyl 3-methylbenzene sulfinate [30P]



3-Methylbenzene disulfide [**29P**] (3.200 g, 12.9 mmol), methanol (65 mL), CH_2Cl_2 (64 mL), and *N*-bromosuccinimide [**46**] (6.889 g, 38.7 mmol) were used as described for **30A**. Column chromatography on silica gel (CH_2Cl_2 eluent) afforded the sulfinate ester, **30P**, as a yellow oil (3.275 g, 74%); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 2.44 (3H, s, C**H**$_3$), 3.49 (3H, s, S-OC**H**$_3$), 7.32-7.39 (1H, m, Ar-**H**), 7.42 (1H, t, *J* = 7.4, Ar-**H**), 7.46-7.56 (2H, m, 2xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 21.5 (CH_3, Aromatic-**C**H$_3$), 49.8 (CH_3, S-O**C**H$_3$), 122.6, 125.8, 129.0, 133.1, (CH, 4xAromatic-**C**H), 139.4 (C, Aromatic-**C**CH$_3$), 144.0 (C, Aromatic-**C**S); m/z (ESI): 171 [(M+H)$^+$], 193 [(M+Na)$^+$]; HRMS (ESI): Exact mass calculated for $C_8H_{11}O_2S$

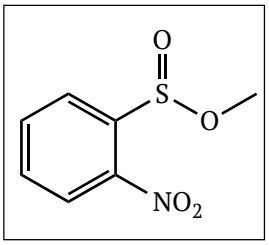
[(M+H)$^+$] 171.0480, Found 171.0483; ν_{max} (ATR)/cm^{-1}: 1124 (S=O), 959 (S-O), 786, 683, 672.

Methyl 4-methylbenzene sulfinate [30Q][5,30]



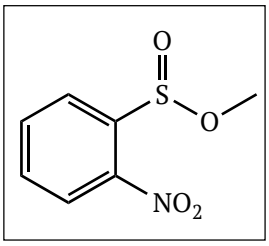
4-Methylbenzene disulfide [**29Q**] (1.565 g, 6.4 mmol), methanol (32 mL), CH_2Cl_2 (32 mL), and *N*-bromosuccinimide [**46**] (3.40 g, 19.1 mmol) were used as described for **30A**. Column chromatography on silica gel (CH_2Cl_2 eluent) afforded the sulfinate ester, **30Q**, as a colourless oil (1.802 g, 83%); ^{1}H NMR: δ_H (300 MHz): 2.43 (3H, s, Ar-C**H$_3$**), 3.47 (3H, s, S-OC**H**$_3$), 7.34 (2H, d, *J* = 8.1, 2xAr-**H**), 7.59 (2H, d, *J* = 8.1, Ar-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 21.6 (CH_3, Aromatic-**C**H$_3$) 49.5 (CH_3, S-O**C**H$_3$), 125.5 (CH, 2xAromatic-**C**H), 129.9 (CH, 2xAromatic-**C**H), 141.1 (C, Aromatic-**C**CH$_3$), 143.0 (C, Aromatic-**C**S); m/z (ESI) 171 [(M+H)$^+$], 193 [(M+Na)$^+$]; ν_{max} (ATR)/cm^{-1}: 1128 (S=O), 957 (S-O), 812, 674, 636, 626. Spectral characteristics are consistent with those reported in the literature.[5,30]

2.8.3.1 First attempted synthesis of methyl 2-nitrobenzene sulfinate [30R][15]



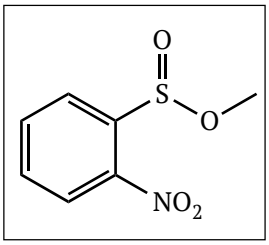
Attempted synthesis of methyl 2-nitrobenzene sulfinate was performed according to the method outlined by Brownbridge and Jowett.[15] 2-Nitrophenyl disulfide [**29R**] (0.269 g, 0.87 mmol), methanol (9 mL), CH_2Cl_2 (9 mL), and *N*-bromosuccinimide [**46**] (0.465 g, 2.61 mmol) were used as described for **30A**. Additional CH_2Cl_2 (20 mL) was required to achieve dissolution of the starting material. An orange colour indicative of reaction progress was not seen for this reaction. TLC monitoring of the reaction showed no disappearance of the disulfide band after 3 h and only starting material was recovered from the reaction mixture.

2.8.3.2 Second attempted synthesis of methyl 2-nitrobenzene sulfinate [30R][15]



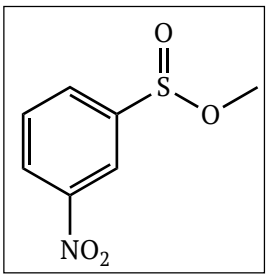
Synthesis of methyl 2-nitrobenzene sulfinate [**30R**] was performed using a modified version of the procedure outlined by Brownbridge and Jowett.[15] 2-Nitrophenyl disulfide [**29R**] (0.423 g, 1.37 mmol), was dissolved in THF (20 mL). The resultant solution was added to a stirred solution of methanol (9 mL) at 0 °C. Upon cooling, additional THF (20 mL) was required to fully dissolve the disulfide product. *N*-Bromosuccinimide [**46**] (0.730 g, 4.10 mmol) was added. A slight orange colouration was observed after approx. 30 min. TLC monitoring of the reaction showed no disappearance of the disulfide band after 3 h and only starting material was recovered from the reaction mixture.

2.8.3.3 Third attempted synthesis of methyl 2-nitrobenzene sulfinate [30R][32]



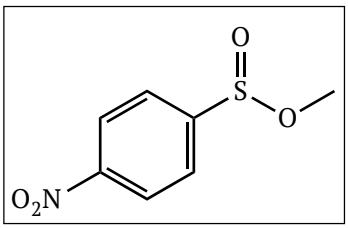
Attempted synthesis of methyl 2-nitrobenzene sulfinate [**30R**] was performed using the method outlined by Meyers and Resek.[32] 2-Nitrophenyl disulfide [**29R**] (0.300 g, 0.97 mmol), was stirred in a solution of methanol (30 mL) and sodium carbonate (0.532 g, 5.0 mmol). The solution was stirred under a nitrogen atmosphere for 5 min and bromine liquid (0.16 mL, 0.48 g, 3.0 mmol) was added dropwise *via* syringe. The reaction was stirred for a period of 3 h, no reaction occurred and the disulfide starting material was recovered from solution.

Methyl 3-nitrobenzene sulfinate [30S][31]



3-Nitrophenyl disulfide [**29S**] (3.108 g, 10.1 mmol), methanol (51 mL), CH_2Cl_2 (50 mL), and *N*-bromosuccinimide [**46**] (5.383 g, 30.2 mmol) were used as described for **30A**. Column chromatography on silica gel (CH_2Cl_2 eluent) afforded the sulfinate ester, **30S**, as an orange oil (3.134 g, 77%); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.58 (3H, s, S-OC**H**$_3$), 7.74-7.83 (1H, m, Ar-**H**), 8.06 (1H, ddd, *J* = 7.7, 1.6, 1.1, Ar-**H**), 8.43 (1H, ddd, *J* = 8.2, 2.3, 1.1, Ar-**H**), 8.54-8.61 (1H, m, Ar-**H**), [Lit.[31] 3.62 (3H, s, S-OC**H**$_3$), 7.33-8.68 (4H, m)]; ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 50.8 (CH_3, S-O**C**H$_3$), 121.0, 127.0, 130.5, 131.3 (CH, 4xAromatic-**C**H), 146.8 (C, Aromatic-**C**S), 148.8 (C, Aromatic-**C**NO_2); m/z (ESI): 202 [$(M+H)^+$]; HRMS (ESI): Exact mass calculated for $C_7H_7O_4NS$ $[M+H]^+$ 202.0175, Found 202.0174; ν_{max} (ATR)/cm^{-1}: 1525 (Aromatic C=C bend), 1348 (N=O), 1130 (S=O) [Lit.[31] 1122], 954 (S-O), 872, 733, 673, 653.

Methyl 4-nitrobenzene sulfinate [30T][14]



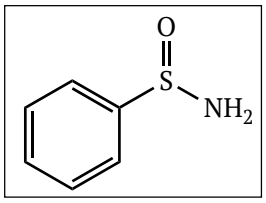
4-Nitrophenyl disulfide [**29T**] (2.000 g, 6.5 mmol), methanol (32 mL), CH_2Cl_2 (64 mL), and *N*-bromosuccinimide [**46**] (3.459 g, 19.4 mmol) were used as described for **30A**. Column chromatography on silica gel (CH_2Cl_2 eluent) afforded the sulfinate ester, **30T**, as an orange solid (1.591 g, 61%); mp 46-48 °C, (Lit.[14] 46-47 °C); ^{1}H NMR: δ_H (300 MHz) ($CDCl_3$): 3.56 (3H, s, S-OC**H**$_3$), 7.88-7.98 (2H, m, 2xAr-**H**), 8.35-8.45 (2H, m, 2xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) ($CDCl_3$): 50.8 (CH_3, S-O**C**H$_3$), 124.4 (CH, 2xAromatic-**C**H), 127.0 (CH, 2xAromatic-**C**H), 150.4 (C, Aromatic-**C**NO_2), 150.5 (C, Aromatic-**C**S); ν_{max} (ATR)/cm^{-1}: 1533 (Aromatic C=C bend), 1344 (N=O), 1313, 1144 (w) (S=O), 1050, 999 (S-O), 849, 744, 722, 681. ^{1}H and ^{13}C NMR characteristics are in agreement with those previously reported.[14]

2.8.4 Synthesis of Substituted Aryl Primary Sulfinamides

General Procedure C:

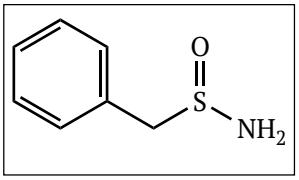
Substituted aromatic sulfinamides were prepared using the methodology outlined by Davis *et al*.[36]

Benzene sulfinamide [31A][37,108]



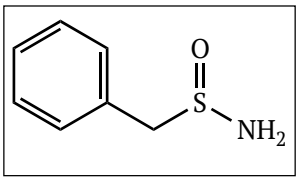
Methyl benzene sulfinate [**30A**] (1.7548 g, 11.2 mmol) was added to a 100 mL round bottom flask, with magnetic stir bar, rubber septum and nitrogen inlet. THF (40 mL) was added to the flask under nitrogen using a syringe. The flask was cooled to approximately -80 °C and lithium bis(trimethyl silyl) amide solution (LiHMDS, **47**), (1.0 M in THF) (14.5 mL, 14.5 mmol) was added via syringe. After 15 min, the flask was removed from the cooling bath and was allowed to warm to room temperature over 45 min with stirring, during which time a cherry-red colour slowly appeared in the flask. The reaction was quenched with saturated ammonium chloride solution (sat. NH_4Cl) (20 mL) and allowed to stir for a further 15 min. Water (20 mL) was added and the reaction mixture was extracted three times with ethyl acetate (20 mL), and the combined organic extracts were washed with saturated ammonium chloride (20 mL), dried (magnesium sulfate) and the solvent removed *in vacuo*. Purification by column chromatography (100% ethyl acetate) afforded the pure sulfinamide, **31A**, as an off white solid (1.137 g, 72%); mp 84-86 °C (Lit.[37] 104-106 °C); ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 6.22 (2H, br s, NH_2), 7.44-7.57 (3H, m, 3xAr-**H**), 7.60-7.69 (2H, m, 2xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) (DMSO-d_6): 125.3 (CH, 2xAromatic-**C**H), 128.6 (CH, 2xAromatic-**C**H), 130.2 (CH, Aromatic-**C**H) 148.2 (C, Aromatic-**C**S); m/z (ESI): 142 [$(M+H)^+$]; ν_{max} (ATR)/cm^{-1}: 3323, 3204 (w) (N-H Stretches), 1520 (Aromatic C=C bend), 1338, 1037 (S=O), 852, 745, 721, 683. Spectral characteristics are in agreement with those previously reported.[37,108]

2.8.4.1 First attempted synthesis of phenylmethanesulfinamide [31B][31]



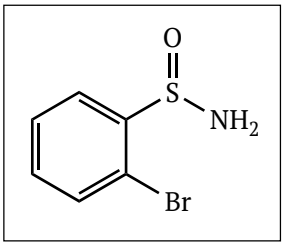
Methyl phenylmethane sulfinate [**30B**] (1.599 g, 9.4 mmol) was added to a 100 mL round bottom flask, with magnetic stirrer, rubber septum and nitrogen inlet. THF (42 mL) was added to the flask under nitrogen using a cannula. The flask was cooled to approx. -80°C and LiHMDS [**47**], (12.7 mL, 12.7 mmol) was added dropwise via syringe. After 15 min, the flask was removed from the cooling bath and was allowed to warm to room temperature with stirring. The reaction was quenched with sat. NH_4Cl (16.5 mL) and allowed to stir for a further 10 min. Water (20 mL) was added and the reaction mixture was extracted three times with ethyl acetate (10 mL), dried (magnesium sulfate) and the solvent removed *in vacuo*. The crude material showed no evidence for formation of the desired product, **31B**, in NMR analysis.

2.8.4.2 Second attempted synthesis of phenylmethanesulfinamide [31B][31]



Methyl phenylmethane sulfinate [**30B**] (0.259 g, 1.5 mmol) was added to a 25 mL round bottom flask, with magnetic stirrer, rubber septum and nitrogen inlet. THF (10 mL) was added to the flask under nitrogen using a syringe. The flask was cooled to approx. -80°C and LiHMDS [**47**], (3.3 mL, 3.3 mmol) was added dropwise via syringe. After 15 min, the flask was removed from the cooling bath and was allowed to warm to room temperature with stirring. The reaction was quenched with sat. NH_4Cl (10 mL) and allowed to stir for a further 10 min. Water (10 mL) was added and the reaction mixture was extracted three times with ethyl acetate (15 mL), dried (magnesium sulfate) and the solvent removed *in vacuo* to yield an oily solid. The crude material showed no evidence of the desired product in NMR analysis.

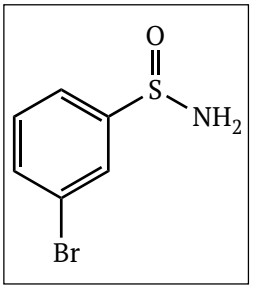
2-Bromobenzene sulfinamide [31C]



Methyl 2-bromobenzene sulfinate [**30C**] (1.9721 g, 8.4 mmol), THF (35 mL), LiHMDS [**47**] (1.0 M in THF, 10.9 mL, 10.9 mmol), and sat. NH_4Cl (15 mL) were used as described for **31A**. Column chromatography on silica gel (40:60 hexane:ethyl acetate eluent) afforded the pure sulfinamide, **31C**, as an off white solid (0.413 g, 18.5%); mp 124-125 °C ; ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 6.24 (2H, br s, NH_2), 7.44 (1H, td, *J* = 7.6, 1.8, Ar-**H**), 7.61 (1H, td, *J* = 7.6, 1.2, Ar-**H**), 7.68 (1H, dd, *J* = 7.9, 1.1, Ar-**H**), 7.92 (1H, dd, *J* = 7.8, 1.7, Ar-**H**); ^{13}C NMR: δ_c (75.5 MHz) (DMSO-d_6): 119.5 (C, Aromatic-**C**Br), 125.5, 128.1, 132.3, 133.0 (CH, 4xAromatic-**C**H) 147.1 (C, Aromatic-**C**S); m/z (ESI): 220 (^{79}Br), 222 (^{81}Br) (1:1) [$(M+H)^+$]; HRMS (ESI): Exact mass calculated for C_6H_7BrNOS (^{79}Br) [$(M+H)^+$] 219.9432, Found 219.9423, C_6H_7BrNOS (^{81}Br) [$(M+H)^+$] 221.9411, Found 221.9401; ν_{max} (ATR)/cm^{-1}: 3323, 3167 (w) (N-H Stretches), 1005 (S=O), 756, 661.

Colourless plate crystals of **31C** were obtained from dissolution of the pure product in dry ethanol and slow evaporation of solvent in a dessicator under nitrogen flow over P_4O_{10} over 5-10 d. Crystal data for **31C**: C_6H_6BrNOS, *Mr* = 220.09, orthorhombic, *Pbca*, *a* = 5.7458(12) Å, *b* = 10.339(2) Å, *c* = 26.279(5) Å, *V* = 1561.1(6) Å^3, *Z* = 8, D_c = 1.873 g cm^{-3}, F_{000} = 864, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 26.79°, μ = 5.461 mm^{-1}, 10552 reflections collected, 1640 unique (R_{int} = 0.0461), final GooF = 1.031, R_1 = 0.0405 [1172 obs. data: $I>2\sigma(I)$], wR_2 = 0.1070 (all data).

3-Bromobenzene sulfinamide [31D]

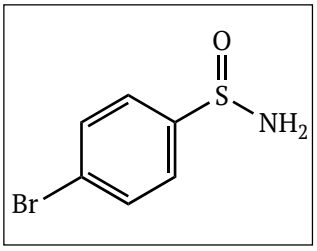


Methyl 3-bromobenzene sulfinate [**30D**] (2.052 g, 8.7 mmol), THF (35 mL), LiHMDS [**47**] (1.0 M in THF, 11.3 mL, 11.3 mmol), and sat. NH_4Cl (15 mL) were used as described for **31A**. The pure sulfinamide, **31D**, was isolated without further purification as a yellow solid (1.664 g, 87%); mp 118-120 °C; ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 6.22 (2H,

br s, NH_2), 7.32 (1H, t, *J* = 7.8, Ar-**H**), 7.43-7.48 (1H, m, Ar-**H**), 7.51-7.56 (1H, m Ar-**H**), 7.59-7.61 (1H, m, Ar-**H**); ^{13}C NMR: δ_c (75.5 MHz) (DMSO-d_6): 121.8 (C, Aromatic-**C**Br), 124.7, 127.8, 130.9, 133.0 (CH, 4xAromatic-**C**H), 150.7 (C, Aromatic-**C**S); m/z (ESI): 220 (^{79}Br), 222 (^{81}Br) (1:1) [$(M+H)^+$]; HRMS (ESI): Exact mass calculated for C_6H_7BrNOS (^{79}Br) [$(M+H)^+$] 219.9432, Found 219.9428, C_6H_7BrNOS (^{81}Br) [$(M+H)^+$] 221.9411, Found 221.9403; ν_{max} (ATR)/cm^{-1}: 3294, 3161 (w) (N-H Stretches), 1017 (S=O), 990, 891, 788, 679.

Pale yellow plate crystals of **31D** were obtained from ethanol over 5-10 d. Crystal data for **31D**: C_6H_6BrNOS, *Mr* = 220.09, orthorhombic, *Pbca*, *a* = 6.141(3) Å, *b* = 7.700(3) Å, *c* = 32.591(14) Å, *V* = 1541.1(12) Å^3, *Z* = 8, D_c = 1.897 g cm^{-3}, F_{000} = 864, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 26.43°, μ = 5.532 mm^{-1}, 17726 reflections collected, 1586 unique (R_{int} = 0.1385), final GooF = 0.999, R_1 = 0.0600 [910 obs. data: $I>2\sigma(I)$], wR_2 = 0.1620 (all data).

4-Bromobenzene sulfinamide [31E][38]

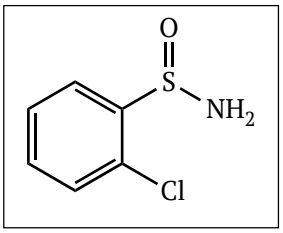


Methyl 4-bromobenzene sulfinate [**30E**] (2.029 g, 8.6 mmol), THF (35 mL), LiHMDS [**47**] (1.0 M in THF, 11.2 mL, 11.2 mmol), and sat. NH_4Cl (15 mL) were used as described for **31A**. Column chromatography on silica gel (40:60 hexane:ethyl acetate eluent) afforded the pure sulfinamide, **31E**, as an off-white solid (0.953 g, 50%); mp 143-145 °C [Lit.[38] 151.2-151.7 °C]; ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 6.34 (2H, br s, NH_2), 7.55-7.64 (2H, m, 2xAr-**H**), 7.70-7.79 (2H, m, 2xAr-**H**), [Lit.[38] ($CDCl_3$) 4.32 (2H, br s), 7.59-7.67 (4H, m)]; ^{13}C NMR: δ_c (75.5 MHz) (DMSO-d_6): 123.8 (C, Aromatic-**C**Br), 127.6 (CH, 2xAromatic-**C**H), 131.5 (CH, 2xAromatic-**C**H), 147.6 (C, Aromatic-**C**S), [Lit.[38] ($CDCl_3$) 126.2, 127.5, 132.3, 145.8]; m/z (ESI): 220 (^{79}Br), 222 (^{81}Br) (1:1) [$(M+H)^+$]; HRMS (ESI): Exact mass calculated for C_6H_7BrNOS (^{79}Br) [$(M+H)^+$] 219.9432, Found 219.9425, C_6H_7BrNOS (^{81}Br) [$(M+H)^+$] 221.9411, Found 221.9403; ν_{max} (ATR)/cm^{-1}: 3270, 3163 (w) (N-H Stretch), 1021 (w) (S=O), 996, 971, 822, 724. Slight differences in the NMR spectra attributed to the use of an

alternative NMR solvent (DMSO-d_6 vs $CDCl_3$), and a higher level of peak resolution in this work.

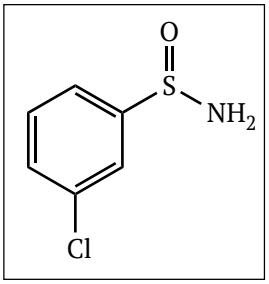
Colourless plates of **31E** were obtained from a mixture of approximately 85:15 $CHCl_3$ and IPA over 7-14 d. Crystal data for **31E**: C_6H_6BrNOS, Mr = 220.09, monoclinic, Pc, a = 18.131(9) Å, b = 5.849(3) Å, c = 7.521(4) Å, β = 92.598(11)°, V = 796.8(7) Å^3, Z = 4, D_c = 1.835 g cm^{-3}, F_{000} = 432, Mo Kα radiation, λ = 0.71073 Å, T = 296(2) K, $2\theta_{max}$ = 25.99°, μ = 5.349 mm^{-1}, 9320 reflections collected, 2760 unique (R_{int} = 0.0795), final GooF = 1.013, R_1 = 0.0975 [1818 obs. data: I>2σ(I)], wR_2 = 0.2681 (all data).

2-Chlorobenzene sulfinamide [31F][28]



Methyl 2-chlorobenzene sulfinate [**30F**] (2.062 g, 10.8 mmol), THF (40 mL), LiHMDS [**47**] (1.0 M in THF, 14.0 mL, 14.0 mmol), and sat. NH_4Cl (20 mL) were used as described for **31A**. The pure sulfinamide, **31F**, was isolated without further purification as a cream solid (1.845 g, 98%); mp 112-114 °C [Lit.[28] 115-117 °C]; ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 6.30 (2H, br s, NH_2), 7.48-7.61 (3H, m, 3xAr-**H**), 7.87-7.95 (1H, m, Ar-**H**); ^{13}C NMR: δ_c (75.5 MHz) (DMSO-d_6): 125.3, 127.6, 129.9 (CH, 3xAromatic-**C**H), 130.4 (C, Aromatic-**C**Cl), 132.2 (CH, Aromatic-**C**H), 145.4 (C, Aromatic-**C**S); m/z (ESI): 176 (^{35}Cl), 178 (^{37}Cl) (3:1) [(M+H)$^+$]; HRMS (ESI): Exact mass calculated for C_6H_7ClNOS (^{35}Cl) [(M+H)$^+$] 175.9937, Found 175.9938; ν_{max} (ATR)/cm^{-1}: 3330, 3218 (w) (N-H Stretches), 1058 (Aromatic C-Cl), 1010 (S=O), 763, 732, 707. Spectral characteristics are in agreement with those reported in the literature.[28]

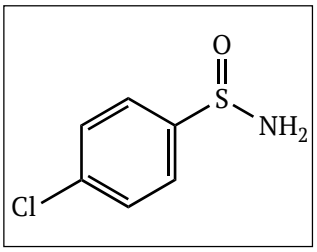
3-Chlorobenzene sulfinamide [31G]



Methyl 3-chlorobenzene sulfinate [**30G**] (1.846 g, 9.7 mmol), THF (40 mL), LiHMDS [**47**] (1.0 M in THF, 12.6 mL, 12.6 mmol), and sat. NH_4Cl (20 mL) were used as described for **31A**. The pure sulfinamide, **31G**, was isolated without further purification as a cream solid (1.545 g, 91%); mp 131-133 °C; ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 6.41 (2H, br s, NH_2), 7.52-7.68 (4H, m, 4xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) (DMSO-d_6): 124.3, 125.0, 130.2, 130.7 (CH, 4xAromatic-**C**H), 133.4 (C, Aromatic-**C**Cl), 150.6 (C, Aromatic-**C**S); m/z (ESI): 176 (^{35}Cl), 178 (^{37}Cl) (3:1) [$(M+H)^+$]; HRMS (ESI): Exact mass calculated for C_7H_8ClNOS (^{35}Cl) [$(M+H)^+$] 175.9937, Found 175.9926, C_7H_8ClNOS (^{37}Cl) [$(M+H)^+$] 177.9907, Found 177.9905; ν_{max} (ATR)/cm^{-1}: 3286 (w) (N-H Stretch), 3061 (w) (Aromatic C-H), 1058 (Aromatic C-Cl), 1014 (S=O), 991, 891, 790, 772, 681, 657.

Colourless brick crystals of **31G** were obtained from ethanol over 5-10 d. Crystal data for **31G**: C_6H_6ClNOS, Mr = 175.63, orthorhombic, *Pbca*, a = 6.1038(16) Å, b = 7.595(2) Å, c = 32.203(8) Å, V = 1492.9(7) Å^3, Z = 8, D_c = 1.563 g cm^{-3}, F_{000} = 720, Mo Kα radiation, λ = 0.71073 Å, T = 296(2) K, $2\theta_{max}$ = 26.45°, μ = 0.715 mm^{-1}, 9971 reflections collected, 1553 unique (R_{int} = 0.0294), final GooF = 1.185, R_1 = 0.0425 [1391 obs. data: I>2σ(I)], wR_2 = 0.1173 (all data).

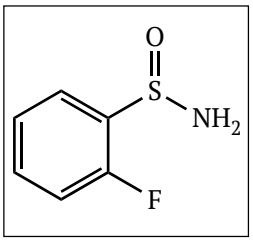
4-Chlorobenzene sulfinamide [31H][37]



Methyl 4-chlorobenzene sulfinate [**30H**] (2.541 g, 13.3 mmol), THF (40 mL), LiHMDS [**47**] (1.0 M in THF, 17.3 mL, 17.3 mmol), and sat. NH_4Cl (20 mL) were used as described for **31A**. The pure sulfinamide, **31H**, was isolated without further purification as a yellow solid (1.926 g, 82%); mp 149-150 (Lit.[37] 130-132 °C); ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 6.34 (2H, br s, NH_2), 7.53-7.71 (4H, m, 4xAr-**H**) [Lit.[37] 4.39 (2H, br s, N**H**$_2$), 7.48 (2H, d, J = 8.7, 2xAr-**H**), 7.68 (2H, d, J = 9.0, 2xAr-**H**)]; ^{13}C NMR: δ_c (75.5

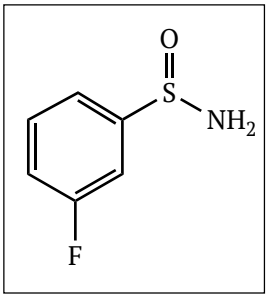
MHz) (DMSO-d_6): 127.4 (CH, 2xAromatic-**C**H), 128.7 (CH, 2xAromatic-**C**H), 135.1 (C, Aromatic-**C**Cl), 147.2 (C, Aromatic-**C**S) [Lit.[37] 127.3, 129.4, 137.8, 145.1]; m/z (ESI): 176 (^{35}Cl), 178 (^{37}Cl) [$(M+H)^+$]; HRMS (ESI): Exact mass calculated for C_6H_7ClNOS (^{35}Cl) $[M+H]^+$, 175.9937 Found 175.9937; ν_{max} (ATR)/cm^{-1}: 3269, 3161 (w) (N-H Stretches), 1079, 1021 (S=O), 826.

2-Fluorobenzene sulfinamide [31I]



Methyl 2-fluorobenzene sulfinate [**30I**] (2.615 g, 15.0 mmol), THF (50 mL), LiHMDS [**47**] (1.0 M in THF, 19.5 mL, 19.5 mmol), and sat. NH_4Cl (25 mL) were used as described for **31A**. The pure sulfinamide, **31I**, was isolated without futher purification as an off-white solid (2.083 g, 87%); mp 90-92 °C; ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 6.37 (2H, br s, NH_2), 7.22-7.45 (2H, m, 2xAr-**H**), 7.49-7.62 (1H, m, Ar-**H**), 7.80 (1H, td, *J* = 7.5, 1.8, Ar-**H**); ^{13}C NMR: δ_c (75.5 MHz) (DMSO-d_6): 116.0 (CH, d, *J* = 20.6, Aromatic-**C**H), 124.7 (CH, d, *J* = 3.5, Aromatic-**C**H), 125.6 (CH, d, *J* = 2.5, Aromatic-**C**H), 132.9 (CH, d, *J* = 7.9, Aromatic-**C**H), 135.1 (C, d, *J* = 15.6, Aromatic-**C**S), 158.0 (C, d, *J* = 247.3, Aromatic-**C**F); m/z (ESI): 160 $[(M+H)^+]$, 182 $[(M+Na)^+]$; HRMS (ESI): Exact mass calculated for C_6H_7FNOS $[(M+H)^+]$ 160.0232, Found 160.0228; ν_{max} (ATR)/cm^{-1}: 3335, 3186 (w) (N-H Stretches), 1469 (Aromatic C=C bend), 1257 (Aromatic C-F), 1215, 1069, 1018 (S=O), 759, 670.

3-Fluorobenzene sulfinamide [31J]

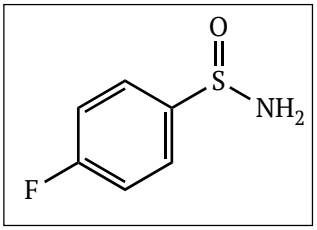


Methyl 3-fluorobenzene sulfinate [**30J**] (1.142 g, 6.5 mmol), THF (25 mL), LiHMDS [**47**] (1.0 M in THF, 8.5 mL, 8.5 mmol), and sat. NH_4Cl (10 mL) were used as described for **31A**. The pure sulfinamide, **31J**, was isolated without further purification as an orange-yellow solid (0.966 g, 94%); mp 138-140 °C; ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 6.37 (2H, br s, NH_2), 7.29-7.53 (3H, m, 3xAr-**H**), 7.54-7.67 (1H, m, Ar-**H**); ^{13}C NMR:

δ_c (75.5 MHz) (DMSO-d_6): 112.7 (CH, d, *J* = 23.8, Aromatic-**C**H), 117.7 (CH, d, *J* = 20.7, Aromatic-**C**H), 122.2 (CH, d, *J* = 2.9, Aromatic-**C**H), 131.3 (CH, d, *J* = 7.8, Aromatic-**C**H), 151.5 (C, d, *J* = 5.4, Aromatic-**C**S), 162.5 (C, d, *J* = 247.4, Aromatic-**C**F); m/z (ESI): 160 [$(M+H)^+$], 182 [$(M+Na)^+$]; HRMS (ESI): Exact mass calculated for C_6H_7FNOS [$(M+H)^+$] 160.0232, Found 160.0227; ν_{max} (ATR)/cm^{-1}: 3269 (N-H Stretch) 3161 (w) (N-H stretch), 3067 (w) (Aromatic C-H), 1466 (Aromatic C=C bend), 1209 (Aromatic C-F), 1015 (S=O), 994, 869, 788, 684, 673.

Yellow plate crystals of **31J** were obtained from ethanol over 5-10 d. Crystal data for **31J**: C_6H_6FNOS, *Mr* = 159.18, orthorhombic, *Pbca*, *a* = 6.0688(12) Å, *b* = 7.5067(15) Å, *c* = 30.438(6) Å, *V* = 1386.7(5) $Å^3$, *Z* = 8, D_c = 1.525 g cm^{-3}, F_{000} = 656, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 23.30°, μ = 0.409 mm^{-1}, 8181 reflections collected, 1230 unique (R_{int} = 0.0533), final GooF = 0.895, R_1 = 0.0383 [884 obs. data: $I>2\sigma(I)$], wR_2 = 0.1413 (all data).

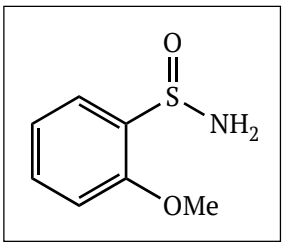
4-Fluorobenzene sulfinamide [31K][28]



Methyl 4-fluorobenzene sulfinate [**30K**] (2.566 g, 14.8 mmol), THF (50 mL), LiHMDS [**47**] (1.0 M in THF, 19.2 mL, 19.2 mmol), and sat. NH_4Cl (25 mL) were used as described for **31A**. The crude sulfinamide was washed with additional NH_4Cl (2 x 25mL) to afford the pure sulfinamide, **31K**, as a pale yellow/cream solid (1.870 g, 80 %); mp 147-149 °C (Lit.[28] 145-147 °C); ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 6.29 (2H, br s, NH_2), 7.30-7.45 (2H, m, 2xAr-**H**), 7.63-7.76 (2H, m, 2xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) (DMSO-d_6): 115.7 (CH, d, *J* = 22.7, 2xAromatic-**C**H), 128.0 (CH, d, *J* = 9.1, 2xAromatic-**C**H), 144.2 (C, d, *J* = 2.7, Aromatic-**C**S), 163.3 (C, d, *J* = 247.0, Aromatic-**C**F); m/z (ESI): 160 [$(M+H)^+$], 182 [$(M+Na)^+$]; ν_{max} (ATR)/cm^{-1}: 3266, 3159 (w) (N-H Stretches), 3067 (w) (Aromatic C-H), 1587 (w) (Aromatic C=C bend), 1479 (w) (Aromatic C=C bend), 1226 (w) (Aromatic C-F), 1019 (S=O), 1004, 834, 814, 614. Spectral characteristics are in agreement with those previously reported.[28]

Yellow plate crystals of **31K** were obtained from ethanol over 5-10 d. Crystal data for **31K**: C_6H_6FNOS, *Mr* = 159.18, orthorhombic, *Pbca*, *a* = 5.9199(16) Å, *b* = 7.522(2) Å, *c* = 30.943(9) Å, *V* = 1377.9(7) Å^3, *Z* = 8, D_c = 1.535 g cm^{-3}, F_{000} = 656, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 25.23°, μ = 0.411 mm^{-1}, 8551 reflections collected, 1238 unique (R_{int} = 0.0777), final GooF = 0.979, R_1 = 0.0513 [885 obs. data: $I>2\sigma(I)$], wR_2 = 0.1631 (all data).

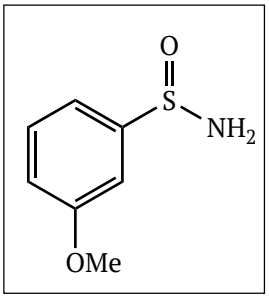
2-Methoxybenzene sulfinamide [31L][14]



Methyl 2-methoxybenzene sulfinate [**30L**] (1.776 g, 9.6 mmol), THF (45 mL), LiHMDS [**47**] (1.0 M in THF, 12.5 mL, 12.5 mmol), and sat. NH_4Cl (25 mL) were used as described for **31A**. Column chromatography on silica gel (ethyl acetate eluent) followed by crystallization from $CHCl_3$ afforded the pure sulfinamide, **31L**, as a white solid (1.180 g, 72%); mp 135-137 °C [Lit.[14] 142-144 °C]; ^{1}H NMR: δ_H (600 MHz) ($CDCl_3$): 3.94 (3H, s, C**H**$_3$), 4.33 (2H, br s, N**H**$_2$), 6.98 (H, d, *J* = 8.2, Ar-**H**), 7.12-7.17 (1H, m, Ar-**H**), 7.46-7.52 (1H, m, Ar-**H**), 7.88 (1H, dd, *J* = 7.6, 1.6, Ar-**H**); ^{13}C NMR: δ_c (150.9 MHz) ($CDCl_3$); 56.1 (CH_3, Ar-O**C**H$_3$), 111.6, 121.2, 125.1, 133.1 (CH, 4xAromatic-**C**H), 134.1 (C, Aromatic-**C**S), 156.3 (C, Aromatic-**C**OCH$_3$); m/z (ESI): 172 [(M+H)$^+$], 194 [(M+Na)$^+$]; ν_{max} (ATR)/cm^{-1}: 3265, 3182 (w) (N-H stretches), 1590 (w) (Aromatic C=C bend), 1477 (Aromatic C=C bend), 1271 (C-O), 1241, 1033 (S=O), 1008, 941, 751, 610. Spectral characteristics are in agreement with those previously reported.[14]

Colourless plate crystals of **31L** were obtained by rapid evaporation from ethanol on a warm clock glass over 1-2 h. Crystal data for **31L**: $C_7H_9NO_2S$, *Mr* = 171.21, hexagonal, P6*cc*, *a* = 19.8271(11) Å, *b* = 19.8271(11) Å, *c* = 7.8368(4) Å, *V* = 2668.0(3) Å^3, *Z* = 12, D_c = 1.279 g cm^{-3}, F_{000} = 1080, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 26.50°, μ = 0.316 mm^{-1}, 1851 reflections collected, 1851 unique (R_{int} = 0.0905), final GooF = 1.097, R_1 = 0.0421 [1601 obs. data: $I>2\sigma(I)$], wR_2 = 0.1219 (all data).

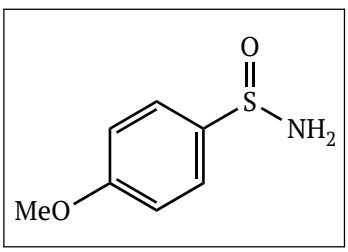
3-Methoxybenzene sulfinamide [31M]



Methyl 3-methoxybenzene sulfinate [**30M**] (1.226 g, 6.8 mmol), THF (35 mL), LiHMDS [**47**] (1.0 M in THF, 8.8 mL, 8.8 mmol), and sat. NH_4Cl (20 mL) were used as described for **31A**. Column chromatography on silica gel (10:90 hexane:ethyl acetate eluent) afforded the pure sulfinamide, **31M**, as an orange solid (0.782 g, 67%); mp 84-86 °C; ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 3.81 (3H, s, C**H$_3$**), 6.24 (2H, br s, N**H$_2$**), 7.03-7.09 (1H, m, Ar-**H**), 7.17-7.25 (2H, m, 2xAr-**H**), 7.41-7.48 (1H, m, Ar-**H**); ^{13}C NMR: δ_c (75.5 MHz) (DMSO-d_6): 55.4 (CH_3, Ar-O**C**H_3), 110.2, 116.2, 117.5, 129.8 (CH, 4xAromatic-**C**H), 149.8 (C, Aromatic-**C**S), 159.4 (C, Aromatic-**C**OCH_3); m/z (ESI): 172 [$(M+H)^+$], 194 [$(M+Na)^+$]; HRMS (ESI): Exact mass calculated for $C_7H_9NO_2S$ $[M+H]^+$, 172.0432 Found 172.0427; ν_{max} (ATR)/cm^{-1}; 3514, 3144 (w) (N-H stretches), 2971 (w) (Alkyl C-H), 1472 (Aromatic C=C bend), 1454 (Aromatic C=C bend), 1424 (Aromatic C=C bend), 1315, 1228 (C-O), 1002 (S=O), 988, 950, 872, 835, 784, 677, 628.

Orange needles of **31M** were used without further recrystallization. Crystal data for **31M**: $C_7H_9NO_2S$, *Mr* = 171.21, tetragonal, *R3c*, *a* = *b* = 22.933(3) Å, *c* = 8.2658(16) Å, γ = 120°, *V* = 3764.8(13) Å^3, *Z* = 18, D_c = 1.359 g cm^{-3}, F_{000} = 1620, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 25.09°, μ = 0.336 mm^{-1}, 16000 reflections collected, 1445 unique (R_{int} = 0.0678), final GooF = 1.031, R_1 = 0.0381 [1094 obs. data: $I>2\sigma(I)$], wR_2 = 0.0843 (all data).

4-Methoxybenzene sulfinamide [31N][37]

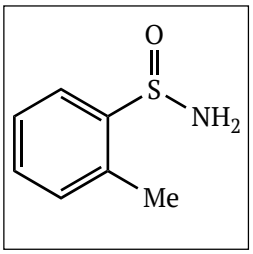


Methyl 4-methoxybenzene sulfinate [**30N**] (1.282 g, 6.9 mmol), THF (30 mL), LiHMDS [**47**] (1.0 M in THF, 9.0 mL, 9.0 mmol), and sat. NH_4Cl (15 mL) were used as described for **31A**. The crude material was recrystallized from ethanol (hexane antisolvent) to give the pure sulfinamide, **31N**, as a cream solid (1.023 g, 87%); mp 128-130 °C

(Lit.[37] 129-131 °C); ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 3.81 (3H, s, C**H**$_3$), 6.12 (2H, br s, N**H**$_2$), 7.02-7.13 (2H, m, 2xAr-**H**), 7.52-7.61 (2H, m, 2xAr-**H**) [Lit.[37] 3.86 (3H, s, Ar-OC**H**$_3$), 4.29 (2H, br s, N**H**$_2$), 7.00 (2H, d, *J* = 9.0, 2xAr_**H**), 7.66 (2H, d, *J* = 9.0, Ar-**H**)]; ^{13}C NMR: δ_c (75.5 MHz) (DMSO-d_6): 55.4 (CH$_3$, Ar-O**C**H$_3$), 114.0 (CH, 2xAromatic-**C**H), 127.0 (CH, 2xAromatic-**C**H), 139.6 (C, Aromatic-**C**S), 160.9 (C, Ar-O**C**CH$_3$); m/z (ESI): 172 [(M+H)$^+$]; ν_{max} (ATR)/cm^{-1}: 3276, 3159 (N-H Stretches), 3067 (Aromatic C-H), 2939 (Alkyl C-H), 1590 (Aromatic C=C bend), 1489 (Aromatic C=C bend), 1471 (Aromatic C=C bend), 1452 (Aromatic C=C bend), 1244 (C-O), 1025 (S=O), 824. ^{13}C NMR data is in agreement with that reported in the literature.[37] Differences in the ^{1}H NMR spectrum can be most likely attributed to the use of an alternative NMR solvent in this work.

Colourless plate crystals of **31N** were obtained from ethanol over 5-10 d. Crystal data for **31N**: $C_7H_9NO_2S$, *Mr* = 171.21, orthorhombic, *Pbca*, *a* = 5.8205(4) Å, *b* = 7.4999(5) Å, *c* = 35.230(3) Å, *V* = 1537.90(18) Å^3, *Z* = 8, D_c = 1.479 g cm^{-3}, F_{000} = 720, Mo Kα radiation, λ = 0.71073 Å, *T* = 100(2) K, $2\theta_{max}$ = 26.42°, μ = 0.366 mm^{-1}, 8065 reflections collected, 1585 unique (R_{int} = 0.0300), final GooF = 1.003, R_1 = 0.0284 [1398 obs. data: $I>2\sigma(I)$], wR_2 = 0.0775 (all data).

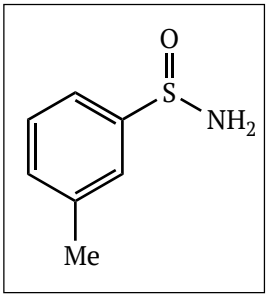
2-Methylbenzene sulfinamide [31O][28]



Methyl 2-methylbenzene sulfinate [**30O**] (1.768 g, 10.4 mmol), THF (40 mL), LiHMDS [47] (1.0 M in THF, 13.5 mL, 13.5 mmol), and sat. NH_4Cl (20 mL) were used as described for **31A**. Column chromatography on silica gel (40:60 hexane:ethyl acetate eluent) afforded the pure sulfinamide**,** **31O**, as a cream solid (0.501 g, 31%); mp 123-126 °C; ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 2.39 (3H, s, Ar-C**H**$_3$), 6.00 (2H, br s, N**H**$_2$), 7.21-7.30 (1H, m, Ar-**H**), 7.35-7.47 (2H, m, 2xAr-**H**), 7.80-7.89 (1H, m, Ar-**H**) [Lit.[28] ($CDCl_3$) 2.47 (3H, s), 4.17 (2H br s), 7.21-7.25 (1H, m), 7.37-7.43 (2H, m), 7.98-8.03 (1H, m)]; ^{13}C NMR: δ_c (75.5 MHz) (DMSO-d_6): 18.2 (CH$_3$, Aromatic-**C**H$_3$), 122.9, 126.1, 130.2, 130.6 (CH, 4xAromatic-**C**H), 135.0 (C, Aromatic-**C**CH$_3$), 145.8 (C, Aromatic-**C**S); m/z (ESI): 156 [(M+H)$^+$]; HRMS (ESI): Exact mass calculated for $C_7H_{10}NOS$ [M+H]$^+$,

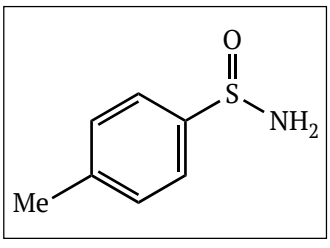
156.0483 Found 156.0481; ν_{max} (ATR)/cm^{-1}: 3271, 3183 (w) (N-H stretches), 3086 (w) (Aromatic C-H), 1563 (w) (Aromatic C=C bend), 1470 (w) (Aromatic C=C bend), 1063, 1018 (S=O), 754, 674. ^{13}C and infrared spectra are in agreement with that reported in the literature.[28]

3-Methylbenzene sulfinamide [31P]



Methyl 3-methylbenzene sulfinate [**30P**] (2.515 g, 14.8 mmol), THF (60 mL), LiHMDS [**47**] (1.0 M in THF, 19.2 mL, 19.2 mmol), and sat. NH_4Cl (30 mL) were used as described for **31A**. The pure sulfinamide, **31P**, was isolated without further purification as a cream solid (2.145 g, 93%); mp 80-83 °C; ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 2.37 (3H, s, C**H**$_3$), 6.19 (2H, br s, N**H**$_2$), 7.28-7.36 (1H, m, Ar-**H**), 7.36-7.49 (3H, m, 3xAr-**H**); ^{13}C NMR: δ_c (75.5 MHz) (DMSO-d_6): 21.0 (CH_3, Ar-C**C**H$_3$), 122.5, 125.6, 128.5, 130.9 (CH, 4xAromatic-**C**H), 138.2 (C, Aromatic-**C**CH$_3$), 148.1 (C, Aromatic-**C**S); m/z (ESI): 156 [(M+H)$^+$]; HRMS (ESI): Exact mass calculated for $C_7H_{10}NOS$ [M+H]$^+$, 156.0483 Found 156.0481; ν_{max} (ATR)/cm^{-1}: 3206 (w) (N-H Stretch), 2992 (w) (Alkyl C-H), 1472 (Aromatic C=C bend), 1449 (Aromatic C=C bend), 1038 (w) (S=O), 995, 952, 778, 693, 601.

4-Methylbenzene sulfinamide [31Q][29,37]

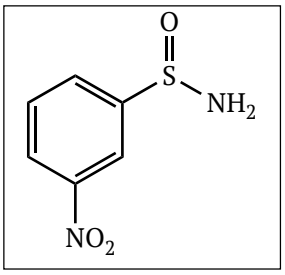


Methyl 4-methylbenzene sulfinate [**30Q**] (1.402 g, 8.2 mmol), THF (35 mL), LiHMDS [**47**] (1.0 M in THF, 10.7 mL, 10.7 mmol), and sat. NH_4Cl (20 mL) were used as described for **31A**. The pure sulfinamide, **31Q**, was isolated without further purification as a white solid (1.026 g, 81%); mp 117-119 °C [Lit.[37] 115-117 °C]; ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 2.36 (3H, s, Ar-C**H**$_3$), 6.16 (2H, br s, NH_2), 7.34 (2H, d, *J* = 8.2, 2xAr-H), 7.54 (2H, d, *J* = 8.2, 2xAr-H) [Lit.[37] 2.42 (3H, s, Ar-C**H**$_3$), 4.30 (2H, br s, N**H**$_2$), 7.30 (2H, d, *J* = 8.1, 2xAr-**H**), 7.62 (2H, d, *J* = 8.1, 2xAr-**H**)]; ^{13}C NMR: δ_c (75.5 MHz)

(DMSO-d_6): 20.8 (CH_3, Aromatic-**C**H_3), 125.4 (CH, 2xAromatic-**C**H), 129.2 (CH, 2xAromatic-**C**H), 140.0 (C, Aromatic-**C**CH_3), 145.2 (C, Aromatic-**C**S) [Lit.[37] 125.6, 129.7, 141.5, 143.6]; m/z (ESI): 156 [$(M+H)^+$]; Exact mass calculated for $C_7H_{10}NOS$ $[M+H]^+$, 156.0483, Found 156.0476; ν_{max} (ATR)/cm^{-1}: 3178 (w) (N-H Stretch), 3092 (w) (Aromatic C-H), 1021 (S=O), 1010, 918, 804, 702. Infrared characteristics are in agreement with that previously reported in the literature for (S_s)-*p*-toluene sulfinamide.[29] Slight differences in the NMR spectra attributed most likely due to the use of an alternative NMR solvent in this work.

Colourless crystals of **31Q** were obtained from ethanol over 5-10 d, unit cell determination was consistent with that found in the literature.[1]

3-Nitrobenzene sulfinamide [31S][31]

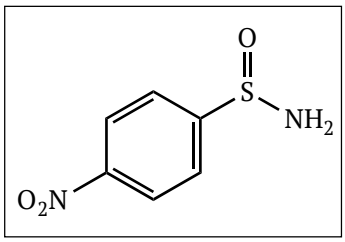


Methyl 3-nitrobenzene sulfinate [**30S**] (2.541 g, 12.7 mmol), THF (50 mL), LiHMDS [**47**] (1.0 M in THF, 16.8 mL, 16.8 mmol), and sat. NH_4Cl (30 mL) were used as described for **31A**. The pure sulfinamide, **31S**, was isolated without further purification as an orange solid (1.758 g, 75%) [Lit.[31] white needles]; mp 107-110 °C (Lit.[31] 110-111 °C); ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 6.37 (2H, br s, NH_2), 7.64 (1H, t, *J* = 7.9, Ar-**H**), 7.83-7.92 (1H, m, Ar-**H**), 8.13-8.20 (1H, m, Ar-**H**), 8.20-8.26 (1H, m, Ar-**H**) [Lit.[31] (CD_3CN) 5.0 (2H, br s, N**H**$_2$), 7.3-8.4 (4H, m)]; ^{13}C NMR: δ_c (75.5 MHz) (DMSO-d_6): 120.1, 125.1, 130.6, 132.2 (CH, 4xAromatic-**C**H), 147.9 (C, Aromatic-**C**NO_2), 150.6 (C, Aromatic-**C**S); m/z (ESI): 187 [$(M+H)^+$]; HRMS (ESI): Exact mass calculated for $C_6H_7N_2O_3S$ $[M+H]^+$, 187.0177 Found 187.0174; ν_{max} (ATR)/cm^{-1}: 3272, 3193 (w) (N-H stretches), 3075 (w) (Aromatic C-H), 1524 (Aromatic C=C bend), 1348 (NO_2), 1065, 1034 (S=O), 999, 810, 692. Slight differences in the NMR spectra attributed to the use of an alternative NMR solvent (DMSO-d_6 vs CD_3CN), and a higher level of peak resolution in this work.

Orange needles of **31S** were obtained from ethanol over 5-10 d. Crystal data for **31S**: $C_6H_6N_2O_3S$, *Mr* = 186.19, orthorhombic, *Pbca*, *a* = 7.8981(6) Å, *b* = 7.5421(6)

Å, c = 28.409(2) Å, V = 1692.3(2) Å^3, Z = 8, D_c = 1.462 g cm^{-3}, F_{000} = 768, Mo Kα radiation, λ = 0.71073 Å, T = 100(2) K, $2\theta_{max}$ = 26.43°, μ = 0.351 mm^{-1}, 9975 reflections collected, 1731 unique (R_{int} = 0.0288), final GooF = 1.029, R_1 = 0.0315 [1549 obs. data: I>2σ(I)], wR_2 = 0.0844 (all data).

4-Nitrobenzene Sulfinamide [31T][14]



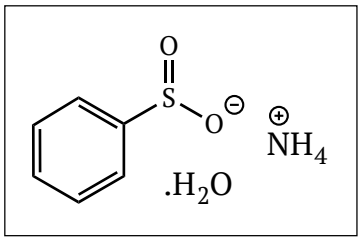
Methyl 4-nitrobenzene sulfinate [**30T**] (1.756 g, 8.7 mmol), THF (45 mL), LiHMDS [**47**] (1.0 M in THF, 11.3 mL, 11.3 mmol), and sat. NH_4Cl (25 mL) were used as described for **31A**. The pure sulfinamide, **31T**, was isolated after column chromatography on silica gel (ethyl acetate eluent) as an orange solid (1.162 g, 72%); mp 148-150 °C (Lit.[14] 158-159 °C); ^{1}H NMR: δ_H (300 MHz) (DMSO-d_6): 6.57 (2H, br s, N**H**$_2$), 7.89-7.95 (2H, m, 2xAr-**H**), 8.34-8.41 (2H, m, 2xAr-**H**) [Lit.[14] 4.45 (2H, br s), 7.96 and 8.37 (4H, AA 'BB' system)]; ^{13}C NMR: δ_c (75.5 MHz) (DMSO-d_6): 123.8 (CH, 2xAromatic-**C**H), 127.1 (CH, 2xAromatic-**C**H), 148.7 (C, Aromatic-**C**NO$_2$), 155.0 (C, Aromatic-**C**S); m/z (ESI): 187 [(M+H)$^+$]; HRMS (ESI): Exact mass calculated for $C_6H_7N_2O_3S$ [M+H]$^+$, 187.0177 Found 187.0171; ν_{max} (ATR)/cm^{-1}: 3324, 3204 (N-H stretches), 3094 (Aromatic C-H), 1520 (Aromatic C=C bend), 1475 (Aromatic C=C bend), 1338 (NO$_2$), 1037 (S=O), 852, 721, 683. ^{13}C NMR is in agreement with that previously reported.[14]

Orange prismatic needles of **31T** were obtained from ethanol over 5-10 d. Crystal data for **31T**: $C_6H_6N_2O_3S$, Mr = 186.19, orthorhombic, $Pna2_1$, a = 25.45(3) Å, b = 6.099(8) Å, c = 5.080(7) Å, V = 788.5(18) Å^3, Z = 4, D_c = 1.568 g cm^{-3}, F_{000} = 384, Mo Kα radiation, λ = 0.71073 Å, T = 296(2) K, $2\theta_{max}$ = 26.52°, μ = 0.376 mm^{-1}, 4581 reflections collected, 1567 unique (R_{int} = 0.0530), final GooF = 0.811, R_1 = 0.0433 [1139 obs. data: I>2σ(I)], wR_2 = 0.1367 (all data).

2.8.5 Synthesis of Ammonium Sulfinate and Sulfonate Salts

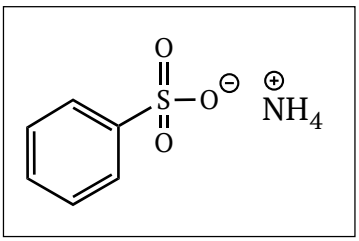
For the purpose of this section, 80% aqueous ethanol refers to a solution of 80% ethanol and 20% water (v/v).

Ammonium benzene sulfinate hydrate [55A]



Benzene sulfinamide [**31A**] (0.1055 g, 0.75 mmol) was stirred in 80% aqueous ethanol for a period of 1 week. The resulting solution was reduced in vacuo to quantitatively yield the desired product, **55A**; mp 192 °C (decomposition); ^{1}H NMR: δ_H (300 MHz) (CD_3OD): 7.34-7.48 (3H, m, 3x Ar-H), 7.58-7.72 (1.8H, m, 2xAr-H); ^{13}C NMR: δ_c (75.5 MHz) (CD_3OD): 125.1 (CH, 2xAromatic **C**-H), 129.5 (CH, 2xAromatic C-H), 130.5 (CH, Aromatic C-H), 157.0 (C, Aromatic **C**-S); m/z (ESI): 141 ($C_6H_4O_2S^-$) [$(M-NH_4)^-$]; HRMS (ESI): Exact mass calculated for $C_6H_5O_2S$ [$M-NH_4]^-$, 141.0010 Found 141.0011; ν_{max} (ATR)/cm^{-1}: 3418 (w), 3353 (w), 3182 (w) (N-H Stretches), 3053 (w) (Aromatic C-H), 1455, 1441 (Aromatic C=C bends), 1036, 1016 (S=O), 956 (S-O), 759, 700; Colourless block crystals of **55A** were obtained by dissolution of benzene sulfinamide [**31A**] in dry ethanol followed by evaporation of solvent to dryness at room temperature under a nitrogen atmosphere.

Crystal data for **55A**: $C_6H_{11}NO_3S$, *Mr* = 177.22, orthorhombic, *Pbca*, *a* = 9.2729(10) Å, *b* = 7.7359(7) Å, *c* = 26.478(3) Å, *V* = 1899.4(3) Å^3, *Z* = 8, D_c = 1.239 g cm^{-3}, F_{000} = 752, Mo Kα radiation, λ = 0.71073 Å, *T* = 300(2) K, $2\theta_{max}$ = 26.37°, μ = 0.305 mm^{-1}, 19082 reflections collected, 1935 unique (R_{int} = 0.0601), final GooF = 1.061, R_1 = 0.0419 [1292 obs. data: $I>2\sigma(I)$], wR_2 = 0.1482 (all data).

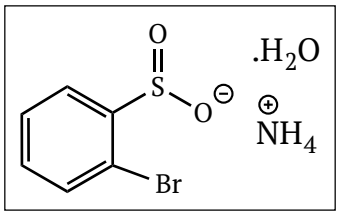


Dissolution of **31A** in bulk methanol followed by evaporation to dryness at room temperature in air resulted in orange plate crystals of **56A.** A pure sample of **56A** could not be obtained for NMR/IR analysis, however the characteristic mass was

identified in HRMS (ESI): Exact mass calculated for $C_6H_5O_3S$ [M-NH_4]$^-$, 156.9959, Found 156.9964.

Crystal data for **56A**: $C_6H_9O_3NS$, *Mr* = 175.20, orthorhombic, *Pbca*, *a* = 7.458(2) Å, *b* = 7.683(2) Å, *c* = 30.490(9) Å, *V* = 1747.1(9) Å^3, *Z* = 8, D_c = 1.332 g cm^{-3}, F_{000} = 736, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 26.36°, μ = 0.331 mm^{-1}, 12205 reflections collected, 1786 unique (R_{int} = 0.0586), final GooF = 1.041, R_1 = 0.0462 [1362 obs. data: $I>2\sigma(I)$], wR_2 = 0.1121 (all data).

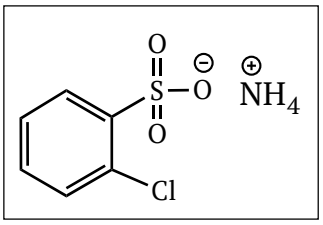
Ammonium 2-bromobenzene sulfinate hydrate [55C]



2-Bromobenzene sulfinamide [**31C**] (0.1031 g, 0.50 mmol) was stirred in 80% aqueous ethanol for a period of 3 weeks. The resulting solution was reduced in vacuo to quantitatively yield the desired product, **55C**; mp 119-121 °C; ^{1}H NMR: δ_H (300 MHz) (CD_3OD): 7.25 (1H, td, *J* = 7.7, 1.6, Ar-**H**), 7.39-7.53 (2H, m, 2xAr-**H**), 7.85 (1H, dd, *J* = 7.7, 1.6, Ar-**H**); ^{13}C NMR: δ_c (75.5 MHz) (CD_3OD): 122.2 (C, Aromatic **C**-Br), 125.5, 129.0, 132.1, 133.7 (CH, 4xAromatic **C**-H), 155.4 (C, Aromatic **C**-S); m/z (ESI): 219 (^{79}Br) ($C_6H_4BrO_2S^-$) [(M-NH_4)$^-$], 221 (^{81}Br) ($C_6H_4BrO_2S^-$) [(M-NH_4)$^-$] (1:0.9); HRMS (ESI): Exact mass calculated for $C_6H_4O_2SBr^-$ (^{79}Br) [(M-NH_4)$^-$]: 218.9115, Found 218.9119, $C_6H_4O_2SBr^-$ (^{81}Br) [(M-NH_4)$^-$]: 220.9095, Found 220.9084; ν_{max} (ATR)/cm^{-1}: 3335 (w), 3232 (w), 3166 (w) (N-H Stretches), 3055 (w) (Aromatic C-H), 1005 (S=O), 956 (S-O), 751, 717.

Colourless plate crystals of **55C** were obtained by dissolution of 2-bromobenzene sulfinamide [**31C**] in acetonitrile followed by evaporation of solvent to dryness at room temperature. Crystal data for **55C**: $C_6H_{10}BrNO_3S$, *Mr* = 256.12, monoclinic, $P2_1/c$, *a* = 14.7557(19) Å, *b* = 4.6846(6) Å, *c* = 14.6873(19) Å, β = 105.095(3)°, *V* = 979.6(2) Å^3, *Z* = 4, D_c = 1.737 g cm^{-3}, F_{000} = 512, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 26.71°, μ = 4.379 mm^{-1}, 14204 reflections collected, 2049 unique (R_{int} = 0.0247), final GooF = 1.056, R_1 = 0.0265 [1826 obs. data: $I>2\sigma(I)$], wR_2 = 0.0733 (all data).

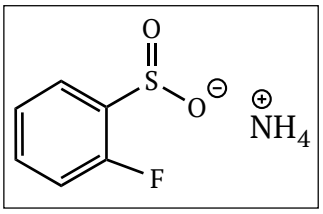
Ammonium 2-chlorobenzenesulfonate [56F]



2-Chlorobenzene sulfinamide [**31F**] (0.1079 g, 0.6 mmol) was stirred in 80% aqueous ethanol for a period of 1 week. The resulting solution was reduced *in vacuo* to quantitatively yield the desired product, **56F**; mp 163-165 °C; ^{1}H NMR: δ_H (300 MHz) (CD_3OD): 7.29-7.44 (3H, m, 3xAr-**H**), 7.81-7.90 (1H, m, Ar-**H**); ^{13}C NMR: δ_c (75.5 MHz) (CD_3OD): 125.2, 128.3, 130.6, 131.8 (CH, 4xAromatic **C**-H), 133.5 (C, Aromatic **C**-Cl), 153.7 (C, Aromatic **C**-S); m/z (ESI): 191 (^{35}Cl) ($C_6H_4SO_2Cl^-$) [$(M-NH_4)^-$], 193 (^{37}Cl) ($C_6H_4SO_2Cl^-$) [$(M-NH_4)^-$] (3:1); HRMS (ESI): Exact mass calculated for $C_6H_4ClO_3S$ (^{35}Cl) [$(M-NH_4)^-$], 190.9570 Found 190.9576, $C_6H_4ClO_3S^-$ (^{37}Cl) [$(M-NH_4)^-$]: 192.9546, Found 192.9548; 3061 (Aromatic C-H), 2854, 1464 (Aromatic C=C bend), 1442 (Aromatic C=C bend), 1393 (S=O), 1009, 962 (S=O),749, 727, 711, ν_{max} (ATR)/cm^{-1}: 1465 (Aromatic C=C bend), 1442 (Aromatic C=C bend), 1009 (S=O), 962 (S-O), 749, 728.

Yellow block crystals of **56F** were obtained by dissolution of 2-chlorobenzene sulfinamide [**31F**] in acetonitrile followed by evaporation of solvent to dryness at room temperature. Crystal data for **56F**: $C_6H_{10}SO_3NCl$, *Mr* = 209.64, monoclinic, $P2_1/c$, a = 8.8685(12) Å, b = 7.5428(10) Å, c = 13.8102(19) Å, β = 92.588(3)°, *V* = 887.6(2) Å^3, *Z* = 4, D_c = 1.569 cm^{-3}, F_{000} = 432, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 26.49°, μ = 0.631 mm^{-1}, 10231 reflections collected, 1831 unique (R_{int} = 0.0329), final GooF = 1.065, R_1 = 0.0310 [1675 obs. data: $I>2\sigma(I)$], wR_2 = 0.0925 (all data).

Ammonium 2-fluorobenzene sulfinate [55I]

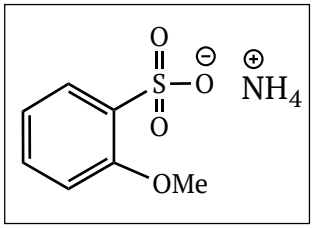


2-Fluorobenzene sulfinamide [**31I**] (0.1017 g, 0.7 mmol) was stirred in 80% aqueous ethanol for a period of 1 week. The resulting solution was reduced in vacuo to quantitatively yield the desired product, **55I**; mp 176-178 °C; ^{1}H NMR: δ_H (300 MHz) (CD_3OD):

7.01-7.11 (1H, m, 2xAr-**H**), 7.23 (1H, td, *J* = 7.5, 1.0, Ar-**H**), 7.32-7.44 (1H, m, Ar-**H**), 7.75 (1H, td, *J* = 7.3, 1.8, Ar-**H**); ^{13}C NMR: δ_c (75.5 MHz) (CD_3OD): 116.5 (CH, d, *J* = 21.6, Aromatic-**C**H), 125.3 (CH, d, *J* = 3.3, Aromatic-**C**H), 125.7 (CH, d, *J* = 4.2, Aromatic-**C**H), 132.4 (CH, d, *J* = 7.5, Aromatic-**C**H), 143.8 (C, d, *J* = 17.8, Aromatic C-S), 161.7 (C, d, *J* = 246.0, Aromatic C-F); m/z (ESI): 159 ($C_6H_4SO_2^-$) [$(M\text{-}NH_4)^-$]; ν_{max} (ATR)/cm^{-1}: 3157 (w) (N-H Stretch), 1464 (Aromatic C=C bend), 1255 (w) (Aromatic C-F), 996 (S=O), 953, 760.

White plate crystals of **55I** were obtained by dissolution of 2-fluorobenzene sulfinamide [**31I**] in ethanol followed by evaporation of solvent to dryness at room temperature. Crystal data for **55I**: $C_6H_8SO_2NF$, *Mr* = 177.19, monoclinic, $P2_1$, a = 10.5287(13) Å, b = 6.7476(8) Å, c = 11.4349(14) Å, *V* = 809.75(17) Å^3, *Z* = 4, D_c = 1.453 g cm^{-3}, F_{000} = 368, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 28.98°, μ = 0.367 mm^{-1}, 7469 reflections collected, 3324 Unique (R_{int} = 0.0588), final GooF = 0.988, R_1 = 0.0465, [1906 obs. Data: $I>2\sigma(I)$]; wR_2 = 0.1285 (all data).

Ammonium 2-methoxybenzenesulfonate [56L]

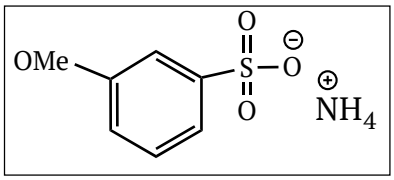


2-Methoxybenzene sulfinamide [**31L**] (0.1201 g, 0.7 mmol) was stirred in 80% aqueous ethanol for a period of 3 weeks with weekly withdrawal of aliquots for periodic NMR analysis in CD_3OD. An analytically pure sample of **56L** could not be obtained, however ^{1}H peaks could be identified in the NMR spectrum, and the characteristic mass was identified; ^{1}H NMR: δ_H (300 MHz) (CD_3OD): 3.89 (3H, s, Ar-CO**H**$_3$), 6.87-7.00 (1H, m, Ar-**H**), 7.07 (1H, d, *J* = 8.3, Ar-**H**), 7.36-7.48 (1H, m, Ar-**H**), 7.84 (1H, dd. *J* = 7.8, 1.8, Ar-**H**); m/z (ESI): 187 ($C_7H_7SO_4^-$) [$(M\text{-}NH_4)^-$]; HRMS (ESI): Exact mass calculated for $C_7H_7O_4S$ [$M\text{-}NH_4]^-$, 187.0065 Found 187.0065.

Colourless needle crystals of **56L** were obtained by dissolution of 2-methoxybenzene sulfinamide [**31L**] in ethanol followed by evaporation of solvent to dryness at room temperature. Crystal data for **56L**: $C_7H_{11}SO_4N$, *Mr* =

205.23, monoclinic, $P2_1/c$, a = 10.7515(16) Å, b = 9.9584(15) Å, c = 8.9522(14) Å, β = 90.498(4)°, V = 958.5(3) Å^3, Z = 4, D_c = 1.422 cm^{-3}, F_{000} = 432, Mo Kα radiation, λ = 0.71073 Å, T = 296.(2) K, $2\theta_{max}$ = 25.29°, μ = 0.321 mm^{-1}, 13144 reflections collected, 3432 unique (R_{int} = 0.0611), final GooF = 0.994, R_1 = 0.0418 [2661 obs. Data: $I>2\sigma(I)$] wR_2 = 0.0910 (all data).

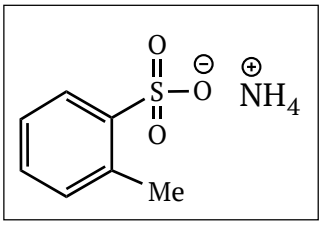
Ammonium 3-methoxybenzene sulfonate [56M]



3-Methoxybenzene sulfinamide [**31M**] (0.0966 g, 0.6 mmol) was stirred in 80% aqueous ethanol for a 3 weeks with weekly withdrawal of aliquots for periodic NMR analysis in CD_3OD. A pure sample of **56M** could not be obtained for NMR/IR analysis**.** However, the characteristic mass was identified; m/z (ESI): 187 ($C_7H_7SO_4^-$) [(M-NH_4)$^-$]; HRMS (ESI): Exact mass calculated for $C_7H_7O_4S$ [M-NH_4]$^-$, 187.0065 Found 187.0061.

Orange plate crystals of **56M** were obtained by dissolution of 3-methoxybenzene sulfinamide [**31M**] in ethanol followed by evaporation of solvent to dryness at room temperature. Crystal data for **56M**: $C_7H_{11}SO_4N$, Mr = 205.23, monoclinic, $P2_1/c$, a = 14.965(4) Å, b = 6.6644(18) Å, c = 9.909(3) Å, β = 99.333(7)°, V = 975.2(5) Å^3, Z = 4, D_c = 1.398 cm^{-3}, F_{000} = 432, Mo Kα radiation, λ = 0.71073 Å, T = 296(2) K, $2\theta_{max}$ = 26.27°, μ = 0.315 mm^{-1}, 13963 reflections collected, 2011 unique (R_{int} = 0.0877), final GooF = 1.027, R_1 = 0.0554 [1523 obs. Data: $I>2\sigma(I)$] wR_2 = 0.1641 (all data).

Ammonium 2-methylbenzenesulfonate [56O]



2-Methylbenzene sulfinamide [**31O**] (0.1076 g, 0.7 mmol) was stirred in 80% aqueous ethanol for a period of 3 weeks with weekly withdrawal of aliquots for periodic NMR analysis in CD_3OD. A pure sample of **56O** could not be obtained for NMR/IR analysis**.** However, the characteristic mass was identified; m/z (ESI): 171 ($C_7H_7SO_3^-$) [(M-

$NH_4)^-$]; HRMS (ESI): Exact mass calculated for $C_7H_7O_3S$ [M-NH_4]$^-$, 171.0116 Found 171.0115.

Yellow plate crystals of **56O** were obtained by dissolution of 2-methylbenzene sulfinamide (**31O**) in ethanol followed by evaporation of solvent to dryness at room temperature. Crystal data for **56O**: $C_7H_{11}SO_3N$, *Mr* = 189.23, monoclinic, $P2_1/c$, a = 15.584(2) Å, b = 5.8532(9) Å, c = 9.9515(16) Å, β = 99.285(4)°, *V* = 895.8(2) Å^3, *Z* = 4, D_c = 1.403 cm^{-3}, F_{000} = 400, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 25.29°, μ = 0.329 mm^{-1}, 6471 reflections collected, 1802 unique (R_{int} = 0.0262), final GooF = 1.025, R_1 = 0.0368 [1276 obs. data: $I>2\sigma(I)$] wR_2 = 0.1003 (all data).

Chapter 2

References

2.9 References

(1) Eccles, K. S.; Morrison, R. E.; Daly, C. A.; O'Mahony, G. E.; Maguire, A. R.; Lawrence, S. E. *CrystEngComm* **2013**, *15*, 7571–7575.

(2) R., G. C.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. *Acta Crystallogr. Sect. B* **2016**, *72*, 171–179.

(3) Davis, F. A.; Reddy, R. E.; Szewczyk, J. M.; Reddy, G. V.; Portonovo, P. S.; Zhang, H.; Fanelli, D.; Thimma Reddy, R.; Zhou, P.; Carroll, P. J. *J. Org. Chem.* **1997**, *62*, 2555–2563.

(4) Liu, G.; Cogan, D. A.; Owens, T. D.; Tang, T. P.; Ellman, J. A. *J. Org. Chem.* **1999**, *64*, 1278–1284.

(5) Davis, F. A.; Reddy, R. E.; Szewczyk, J. M. *J. Org. Chem.* **1995**, *60*, 7037–7039.

(6) Garcia Ruano, J. L.; Aleman, J.; Belen Cid, M.; Parra, A. *Org. Lett.* **2005**, *7*, 179–182.

(7) Liu, G.; Cogan, D. A.; Ellman, J. A.; June, R. V. *J. Am. Chem. Soc.* **1997**, *119*, 9913–9914.

(8) Bolm, C.; Bienewald, F. *Agnew. Chem. Int. Ed. Engl.* **1995**, *34*, 2640–2642.

(9) Robak, M. T.; Herbage, M. A.; Ellman, J. A. *Chem. Rev.* **2010**, *110*, 3600–3740.

(10) Furukawa, M.; Okawara, T. *Synthesis.* **1976**, *5*, 339–430.

(11) Harmata, M.; Zheng, P.; Huang, C.; Gomes, M. G.; Ying, W.; Ranyanil, K.; Balan, G.; Calkins, N. L. *J. Org. Chem.* **2007**, *72* (11), 683–685.

(12) Barry, N.; Lawrence, S. E.; Maguire, A. R. *Unpublished Work*.

(13) Fulton, J. R.; Kamara, L. M.; Morton, S. C.; Rowlands, G. J. *Tetrahedron* **2009**, *65*, 9134–9141.

(14) Garcia Ruano, J. L.; Parra, A.; Alemán, J.; Fajardo, C. *Org. Lett.* **2005**, *7*, 5493–5496.

(15) Brownbridge, P.; Jowett, I. C. *Synthesis.* **1987**, 252–254.

(16) Asai, Y.; Ogawa, S.; Noguchi, T.; Hatano, A.; Hirai, Y.; Kirihara, M.; Asai, Y.; Ogawa, S.; Noguchi, T.; Hatano, A.; Hirai, Y. *Synthesis.* **2007**, *21*, 3286–3289.

(17) Chai, P. J.; Li, Y. S.; Tan, C. X. *Chinese Chem. Lett.* **2011**, *22*, 1403–1406.

(18) Banfield, S. C.; Omori, A. T.; Leisch, H.; Hudlicky, T. *J. Org. Chem.* **2007**, *72*, 4989–4992.

(19) Vandavasi, J. K.; Hu, W. P.; Chen, C. Y.; Wang, J. J. *Tetrahedron* **2011**, *67*, 8895–8901.

(20) Zhao, J.; Song, T.; Jiang, H.; Xu, L.; Zhu, S. *Chinese J. Chem.* **2010**, *28*, 1623–1629.

(21) García Ruano, J. L.; Parra, A.; Alemán, J. *Green Chem.* **2008**, *10*, 706–711.

(22) Taniguchi, N. *Synlett* **2005**, *11*, 1687–1690.

(23) Figuly, G. D.; Martin, J. C. *J. Org. Chem.* **1980**, *45*, 3728–3729.

(24) Trofimov, B. A.; Sinegovskaya, L. M.; Gusarova, N. K. *J. Sulfur Chem.* **2009**, *30*, 518–554.

(25) Dauben Jr., H. J.; McCoy, L. L. *J. Am. Chem. Soc.* **1959**, *81*, 4863.

(26) Martin, C.; Sandrinelli, F.; Perrio, C.; Perrio, S.; Lasne, M. C. *J. Org. Chem.* **2006**, *71*, 210–214.

(27) Du, B.; Li, Z.; Qian, P.; Han, J.; Pan, Y. *Chem. - An Asian J.* **2016**, *11* (4), 478–481.

(28) Kowalczyk, R.; Edmunds, A. J. F.; Hall, R. G.; Bolm, C. *Prep. Chromatogr.* **2011**, *13*, 1–156.

(29) Zhu, R. H.; Shi, X. X. *Tetrahedron Asymmetry* **2011**, *22*, 387–393.

(30) Rayner, P. J.; Brien, P. O.; Horan, R. A. J. *J. Am. Chem. Soc.* **2013**, *135*, 8071–8077.

(31) Davis, F. A.; Friedman, A. J.; Nadir, U. K. *J. Am. Chem. Soc.* **1978**, *100*, 2844–2852.

(32) Meyers, A. I.; Resek, J. E. *Tetrahedron Lett.* **1995**, *36*, 7051–7054.

(33) Kucsman, A.; Kapovits, I. *J. Mol. Struct.* **1989**, *198*, 339–353.

(34) Luxon, S. G. *Hazards in the Chemical Laboratory*, 5th ed.; Luxon, S. G., Ed.; Royal Society of Chemistry, 1992.

(35) Subasinghe, N. L.; Ballentine, S.; Travins, J. M.; Khalil, E. M.; Ali, F.; Leonard, K. A.; Gushue, J. M.; Winters, M. P.; Hufnagel, H.; Cummings, M. D. Novel Thiophene Sulfoximines for treating complement-mediated diseases and conditions. WO2006/101860, 2006.

(36) Davis, F. A.; Zhang, Y.; Andemichael, Y.; Fang, T.; Fanelli, D. L.; Zhang, H.; R, S.; Soc, M. J. C. *J. Org. Chem.* **1999**, *64*, 1403–1406.

(37) Savile, C. K.; Magloire, V. P.; Kazlauskas, R. J. *J. Am. Chem. Soc.* **2005**, *127*, 2104–2113.

(38) Pochetti, G.; Gavuzzo, E.; Campestre, C.; Agamennone, M.; Tortorella, P.; Consalvi, V.; Gallina, C.; Hiller, O.; Tschesche, H.; Tucker, P. A.; Mazza, F. *J. Med. Chem.* **2006**, *49*, 923–931.

(39) Erlenmeyer, H.; Seiler, H. *Helv. Chim. Acta* **1957**, *40*, 88–89.

(40) Hou, X.; Li, B.; Zhang, M.; Dai, L. 1-Substituted-N-sulfinyl-2-aziridine methylenimines preparation and usage, Patent, August 15, 2001.

(41) Hunter, C. A. *Angew. Chemie - Int. Ed.* **2004**, *43*, 5310–5324.

(42) Shan, N.; Zaworotko, M. J. *Drug Discov. Today* **2008**, *13*, 440–446.

(43) Qiao, N.; Li, M.; Schlindwein, W.; Malek, N.; Davies, A.; Trappitt, G. *Int. J. Pharm.* **2011**, *419*, 1–11.

(44) Aakeroy, C. B.; Beatty, A. M.; Helfrich, B. A. *J. Am. Chem. Soc.* **2002**, *124*, 14425–14432.

(45) Macrae, C. F.; Bruno, I. J.; Chisholm, J. A.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; van de Streek, J.; Wood, P. A. J. *J. Appl. Crystallogr.* **2008**, *41*, 466–470.

(46) Leiserowitz, L.; Schmidt, G. M. J. *J. Chem. Soc.* **1969**, *500*, 2372–2382.

(47) Taylor, R.; Allen, F. H.; Cole, J. C. *CrystEngComm* **2015**, *17*, 2651–2666.

(48) Leiserowitz, L.; Tuval, M. *Acta Crystallogr. Sect. B* **1978**, *34*, 1230–1247.

(49) Thun, J.; Seyfarth, L.; Senker, J.; Dinnebier, R. E.; Breu, J. *Angew. Chemie - Int. Ed.* **2007**, *46*, 6729–6731.

(50) Blake, C. C. F.; Small, R. W. H. *Acta Crystallogr. Sect. B* **1972**, *28*, 2201–2206.

(51) Macdonald, C.; Whitesides, M. *Chem. Rev.* **1994**, *94*, 2383–2420.

(52) Aakeröy, C. B.; Scott, B. M. T.; Desper, J. *New J. Chem.* **2007**, *31*, 2044–2051.

(53) Eccles, K. S.; Elcoate, C. J.; Maguire, A. R.; Lawrence, S. E. *Cryst. Growth Des.* **2011**, *11*, 4433–4439.

(54) Jan, C. Y.; Shamsudin, N. B. H.; Tan, A. L.; Young, D. J.; Ng, S. W.; Tiekink, E. R. T. *Acta Crystallogr. Sect. E Struct. Reports Online* **2014**, *70*, 293.

(55) Rheingold, A. L.; Baldacchini, C. J.; Grote, C. W. *J. Crystallogr. Spectrosc. Res.* **1989**, *19*, 25–37.

(56) Izumi, T.; Okamoto, N. *Mem. Chubu Inst. Technol.* **1972**, *8*, 139.

(57) Kato, Y.; Taniguchi, T.; Takaki, Y.; Sakata, K. *Mem. Osaka Kyoiku Univ. Ser. 3* **1967**, *16*, 45.

(58) Tothadi, S.; Joseph, S.; Desiraju, G. R. *Cryst. Growth Des.* **2013**, *13*, 3242–3254.

(59) Kato, Y.; Takaki, Y.; Sakurai, K. *Acta Crystallogr. Sect. B* **1974**, *30*, 2683–2687.

(60) Hattori, S.; Taniguchi, T.; Sakurai, K. *Mem. Osaka Kyoiku Univ. Ser. 3* **1975**, *24*, 35.

(61) Hayashi, T.; Nakata, K.; Takaki, Y.; Sakurai, K. *Bull. Chem. Soc. Japan* **1980**, *53*, 801.

(62) Taniguchi, B. Y. T.; Nakata, K.; Takaki, Y.; Sakurai, K. *Acta Crystallogr. Sect. B* **1978**, *34*, 2574–2578.

(63) Kato, Y.; Sakurai, K. *Bull. Chem. Soc. Japan* **1982**, *55*, 1643.

(64) Taniguchi, T.; Takaki, Y.; Sakurai, K. *Mem. Osaka Kyoiku Univ. Ser. 3* **1975**, *24*, 119.

(65) Kubota, M.; Ohba, S. *Acta Crystallogr. Sect. B* **1992**, *48*, 849–854.

(66) Hathwar, V. R.; Row, T. N. G. *Cryst. Growth Des.* **2011**, *11*, 1338–1346.

(67) Moribe, K.; Tsuchiya, M.; Tozuka, Y.; Yamaguchi, K.; Oguchi, T.; Yamamoto, K. *J. Incl. Phenom. Macrocycl. Chem.* **2006**, *54*, 9–16.

(68) Kato, Y.; Yamazaki, M.; Yokota, M. *Mem. Osaka Kyoiku Univ. Ser. 3* **1979**, *27*, 125.

(69) Orii, S.; Nakamura, T.; Takaki, Y.; Y, S.; Kakudo, M. *Bull. Chem. Soc. Japan* **1963**, *36*, 788–793.

(70) Kato, Y.; Oguru, N.; Yagi, K. *Mem. Osaka Kyoiku Univ. Ser. 3* **1981**, *29*, 91.

(71) Fujimori, K.; Tsukihara, T.; Katsube, Y.; Yamamoto, J. *Bull. Chem. Soc. Japan* **1972**, *45*, 1564.

(72) Nakata, K. *Mem. Osaka Kyoiku Univ. Ser. 3* **1987**, *36*, 15.

(73) Tonogaki, M.; Kawata, T.; Ohba, S.; Iwata, Y.; Shibuya, I. *Acta Crystallogr. Sect. B* **1993**, *49* (6), 1031–1039.

(74) Bhatt, R.; Vries, P. De; Tulinsky, J.; Bellamy, G.; Baker, B.; Singer, J. W.; Klein, P. *J. Med. Chem.* **2003**, *20*, 121–128.

(75) Revankar, G. R.; Hanna, N. B.; Ramasamy, K.; Larson, S. B.; Smee, D. F.;

Finch, R. A.; Avery, T. L.; Robins, R. K. *J. Heterocylic Chem.* **1990**, *27*, 909.

(76) Keceli, G.; Toscano, J. P. *Biochemistry* **2012**, *51*, 4206–4216.

(77) Etter, M. C. *J. Phys. Chem.* **1991**, *95*, 4601–4610.

(78) Mahieux, J.; Sanselme, M.; Coquerel, G. *Cryst. Growth Des.* **2016**, *16*, 396–405.

(79) Taylor, R. *Cryst. Growth Des.* **2016**, acs.cgd.6b00736.

(80) Daszkiewicz, M. *CrystEngComm* **2013**, *15*, 10427–10432.

(81) Spek, A. L. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2009**, *65*, 148–155.

(82) Biasotti, J. B.; Andersen, K. K. *J. Am. Chem. Soc.* **1971**, *93*, 1178–1182.

(83) Asefi, H.; Tilett, J. G. *J. Chem. Soc. Perkin Trans. 2* **1979**, 1579–1582.

(84) Wagner, B. J.; Takahashi Doi, J.; Musker, W. K. *J. Org. Chem.* **1990**, *55*, 5940–5945.

(85) Bagno, A.; Eustace, S. J.; Johansson, L.; Scorrano, G. *J. Org. Chem.* **1994**, *59* (1), 232–233.

(86) Bujnicki, B.; Kolbe, A.; Stefaniak, L. *J. Org. Chem.* **1996**, *3263*, 7593–7596.

(87) Kyung Kim, C.; Lee, I. *Bull. Korean Chem. Soc.* **1997**, *18*, 880–885.

(88) Okuyama, T.; Nagase, S. *J. Chem. Soc. Perkin Trans.* **1994**, 1011–1014.

(89) Piggott, A. M.; Karuso, P. *Tetrahedron Lett.* **2007**, *48*, 7452–7455.

(90) Keceli, G.; Moore, C. D.; Labonte, J. W.; Toscano, J. P. *Biochemistry* **2013**, *52*, 7387–7396.

(91) Bujnicki, B.; Drabowicz, J.; Mikolajczyk, M. *Molecules* **2015**, *20*, 2949–2972.

(92) Okaya, Y. *Acta Crystallogr.* **1967**, *22*, 104–110.

(93) Rogers, R. D.; Bond, A. H.; Henry, R. F. *Acta Crystallogr. Sect. C Cryst. Struct. Commun.* **1991**, *47*, 168–170.

(94) Schreuer, J. *Z. Krist. New Cryst. Struct.* **1999**, *214*, 315.

(95) Fewings, K. R.; Junk, P. C.; Georganopoulou, D.; Prince, P. D.; Steed, J. W. *Polyhedron* **2001**, *20*, 643–649.

(96) Wang, K. W.; Feng, W. J.; Li, H. Y.; Ma, L. L.; Jin, Z. M. *Acta Crystallogr. Sect. E Struct. Reports Online* **2007**, *63*, 3481.

(97) Barzilovich, P. Y.; Lyssenko, K. A.; Antipin, M. Y.; Aldoshin, S. M. *Russ. Chem. Bull.* **2011**, 1159.

(98) Suarez, S.; Doctorovich, F.; Baggio, R. *Acta Crystallogr. Sect. E Struct.*

Reports Online **2012**, *68*, 2228–2229.

(99) Jansch, D.; Schollmeyer, D. *CSD Priv. Commun. DUTZEX03* **2014**, CCDC 1026339.

(100) Zheng, X. F.; Zhu, L. G. *J. Mol. Struct.* **2014**, *1065–1066*, 113–119.

(101) Perrin, D. D. *Purification of Laboratory Reagents*; Pergamon Press, 1996.

(102) Eccles, K. S.; Stokes, S. P.; Daly, C. A.; Barry, N. M.; McSweeney, S. P.; O'Neill, D. J.; Kelly, D. M.; Jennings, W. B.; Ni Dhubhghaill, O. M.; Moynihan, H. A.; Maguire, A. R.; Lawrence, S. E. *J. Appl. Crystallogr.* **2011**, *44*, 213–215.

(103) Bruker AXS: Madison, WI 2013.

(104) Sheldrick, G. M. *Acta Crystallogr. Sect. A* **2008**, *64*, 112–122.

(105) Wave Function, Inc., Irvine, CA.

(106) Wolff, S. K.; Grimwood, D. J.; McKinnon, J. J.; Turner, M. J.; Jayatilaka, D.; Spackman, M. A. Perth, Australia 2012.

(107) Barba, F.; Ranz, F.; Batanero, B. *Tetrahedron Lett.* **2009**, *50*, 6798–6799.

(108) Funes Maldonado, M.; Sehgelmeble, F.; Bjarnemark, F.; Svensson, M.; Ahman, J.; Arvidsson, P. I. *Tetrahedron* **2012**, *68*, 7456–7462.

Chapter 3

Cocrystallization of Salsalate

3.1 Introduction to Salsalate

The salicylate class of anti-inflammatory drugs have been used in clinical practice for over 250 years. In fact, salicylates were first referred to in the 5th century BC by Hippocrates as 'the bitter powder extract from willow bark that eased aches and pains and reduced fevers'.[1] The recognition of the efficacy of salicylic acid (as extracted from willow bark) is credited to the Reverend Edward Stone in 1763.[2] Aspirin [**57**] [acetylsalicylic acid] was synthesised in 1897 by Hoffman, and has since become 'the most popular painkiller in the world'.[2] Salsalate [**58**], the generic name for 2-(2-hydroxybenzoyl)oxybenzoic acid, is a non-steroidal anti-inflammatory drug (NSAID) of the salicylate class. It is derived from the self condensation of salicylic acid and is the pro-drug of salicylate [**59**] – the metabolite which confers pharmaceutical effect (Figure 1).

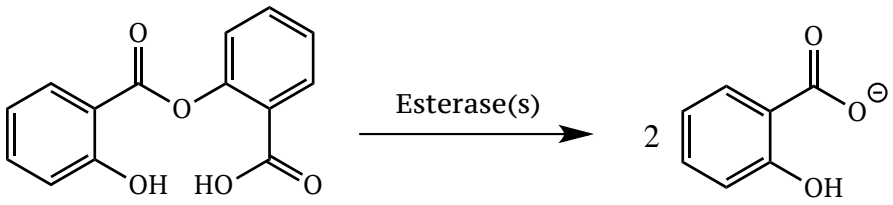


Figure 1: Metabolism of 58 pro-drug to the active form, salicylate [59].

58 is marketed under several proprietary names, including Disalcid[3], Diplosal,[4] Disalgesic,[3] Umbradol,[3] and Mono-Gesic.[3] The primary use of this material is the treatment of inflammation and pain caused by rheumatoid arthritis, osteoarthritis and other rheumatological conditions.[1] **58** has been shown to have near-equivalent molar efficacy to aspirin in the treatment of pain and soft-tissue inflammation, while also providing a low risk of upper gastrointestinal ulcers.[4] Gastrointestinal (GI) injury is a common issue associated with NSAID use. Many NSAIDs (aspirin and naproxen in particular) impair production of prostaglandins and block thromboxane production. Prostaglandins act to protect the GI muscosal membranes, and thromboxane acts to clot an active site of bleeding. This means that extended use of these NSAID products can result in GI injury by reducing protection of the GI mucosal membrane and exacerbating any active site of bleeding present in the GI tract.[5] In contrast, comparative tests of **57** vs **58** have concluded that the occurrence of GI irritations are significantly lowered

with the use of **58**. In one such study, of the 10 patients treated with each drug, only 1 experienced mild gastric irritation when treated with **58**, versus 6 out of 10 experiencing moderate to severe irritation with **57** treatment.[6]

58 is poorly water soluble, and insoluble in acidic media (stomach acid) meaning that it bypasses gastric absorption and so has significantly reduced effect on important prostaglandins, leading to reduced GI damage.[6] This means that **58** is an attractive option for patients that require extended NSAID treatment. The rate of metabolism of **58** to **59** determines its efficacy as a treatment.[4] A significant portion of the API passes through the body unchanged; up to 13% of the ingested medication is excreted as the glucoronide conjugate of the parent material *via* the kidneys.[7] This means that less of the active form of the compound is available for pharmaceutical effect.

58 has recently been investigated for the treatment of diabetes, because it leads to insulin sensitisation and glycemic improvement (reduced blood-glucose levels), with little side-effects;[8] and has shown promise in reducing tau levels *via* inhibition of acetyltransferase p300-induced tau acetylation in mice. Tau levels are associated with progression of neurodegenerative diseases such as Alzheimer's disease and frontotemporal dementia.[9] This research is opening new avenues for the full potential of salsalate as a treatment, and since salsalate is administered as a solid material, creating a requirement for a full understanding of the behaviour of this material in the solid state.

3.1.1 Salsalate in the Solid State

A search of the CSD[10] yields two crystal structures for **58**, WOQDAH[11] (monoclinic space group *Cc*) and WOQDAH01[3] (orthorhombic space group *Fdd2*). However, both the reduced and the Niggli cells of the two crystal structures are remarkably similar (Table 1).

Table 1: Niggli cell and reduced cell values for WOQDAH and WOQDAH01[3,11]

Value	WOQDAH[11]		WOQDAH01[3]	
	Niggli Cell	Reduced Cell	Niggli Cell	Reduced Cell
a	83.16	9.17 Å	83.86	9.16 Å
b	83.16	9.17 Å	83.86	9.16 Å
c	241.18	15.53 Å	242.2	15.57 Å
α	-41.01	106.7°	-41.87	107.0°
β	-41.01	106.7°	-41.87	107.0°
γ	-0.11	90.1°	-0.30	90.0°

Both crystal structures of **58** show disorder created by atropisomerism around the ester carbonyl-phenyl bond (C_7-C_8, Figure 2). This leads to two competing intramolecular interactions of the phenolic O-H in the molecule, both of which are S(6). The first conformer occurs with the phenol donating its hydrogen bond to the ester carbonyl; this is the major conformer for the molecule, apparent in 72% of cases. The minor conformer (28%) includes the phenol in a bifurcated hydrogen bond to the acid carbonyl and the ester oxygen.[3] The major configuration (Figure 2, left) is associated with the strongest intramolecular hydrogen bond.[3] However, both configurations create stable intramolecular six-membered rings in agreement with Etter's rules (Section 1.1.2).[12]

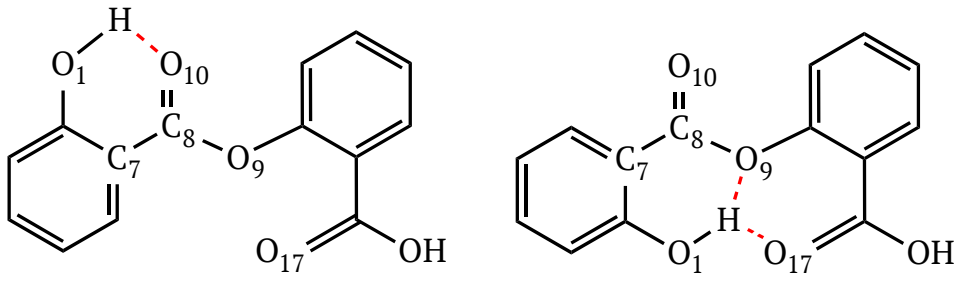


Figure 2: Intramolecular six-membered ring formed in the major (left) and minor (right) conformers of 58 [hydrogen bonds shown in red].[3]

The primary intermolecular interaction in **58** is the $R_2^2(8)$ dimer formed between the carboxylic acid functional groups on two separate molecules (Figure 3); with the carboxylic acid functional groups offset by 33° (Figure 4).

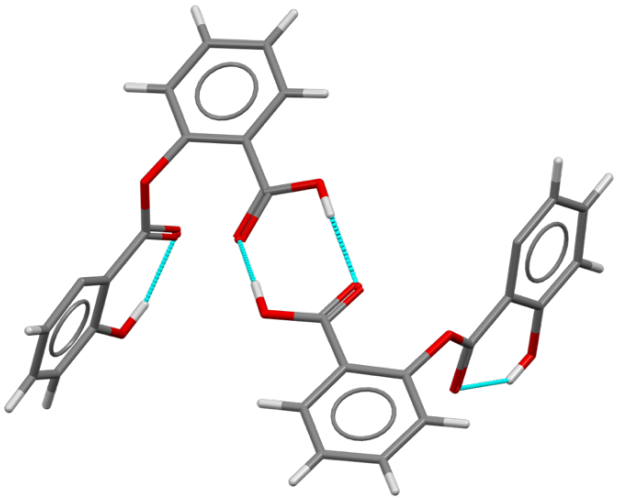


Figure 3: The $R_2^2(8)$ dimer motif present in 58 (WOQDAH) [major conformer shown].[11]

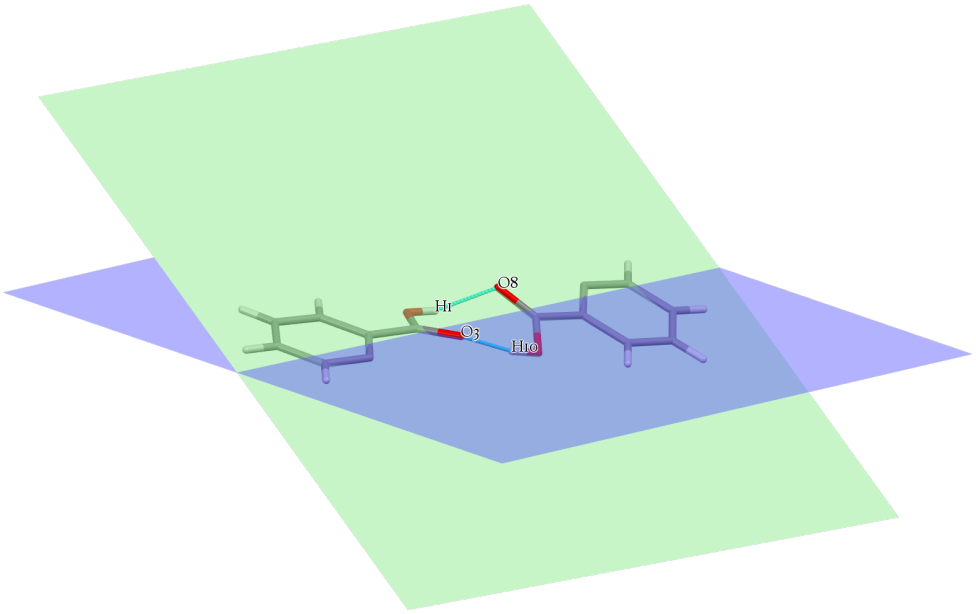


Figure 4: The offset $R_2^2(8)$ dimer motif of 58 (WOQDAH) [major conformer shown, and phenyl ester removed for clarity].[11]

There are no other identifiable hydrogen bonds in this structure, with crystal growth achieved by translation about a 2-fold rotation axis (Figure 5). However, Mercury has identified a very weak H-O···O=C interaction at a distance of 3.0 Å creating a chain down the *c*-axis. The long distance of the interaction suggests that it is more likely a result of close packing, and not a true hydrogen bond.

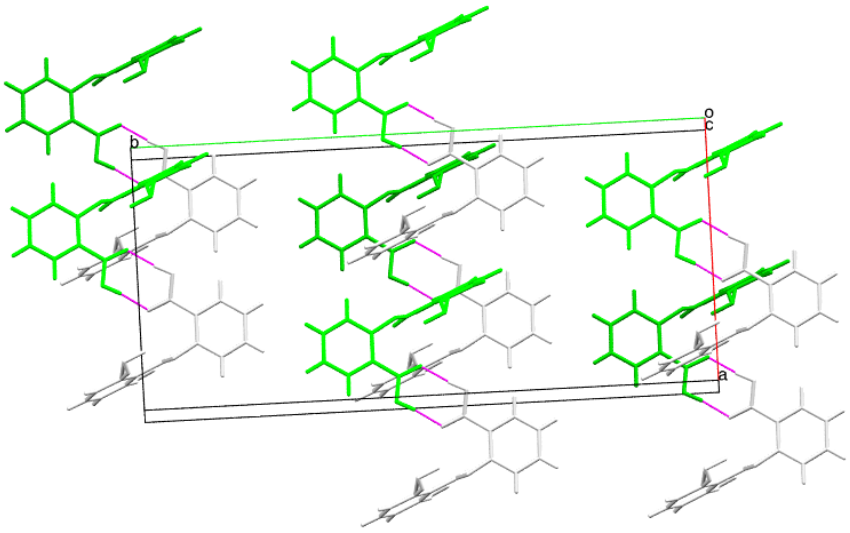


Figure 5: Packing diagram of WOQDAH01 showing the dimer pairs related by 2-fold rotation about the *c*-axis [hydrogen bonds in magenta]

During this work, attempts were made to determine a stable crystal structure of Form I without the rotational disorder observed in both WOQDAH and WOQDAH01. However, in all attempts, this atropisomerism around C_7-C_8 persisted in the crystal structure. It seems that there is little energetic barrier to the rotation around this bond, and the formation of the S(6) in both orientations renders the positions almost equivalent.

3.1.2 Project Objectives

The current CSD[10] does not contain any polymorphs or cocrystals of **58**. This gap in the current solid state literature for this material provides strong rationale for an investigation into novel solid forms of **58**, leading to the following objectives for the project:

(a) Investigate polymorphism of **58** from a range of solvents.

(b) Perform a cocrystallization screen using traditional solid state screening techniques to identify novel solid forms of **58** using knowledge-based coformer selection methods.

(c) Characterise any novel solid forms of **58** identified.

3.2 Polymorphic Screen

As shown above only one, slightly disordered, crystal structure of **58** is available in the literature, and so it was considered prudent to conduct a polymorphic screen to investigate if another stable or metastable crystal structure existed for the material. Methanol, ethanol, acetone, ethyl acetate, acetonitrile, hexane, tetrahydrofuran, dichloromethane, chloroform and water were used for testing. Approximately 0.05 g of **58** was added to a 12 mm glass sample vial, with approximately 5 mL of the relevant solvent. Each sample was capped and a pinhole added to allow complete evaporation of the solvent over 2-3 weeks. Screening was performed using IR and SCXRD (Table 2).

Table 2: Results of polymorphic screen

Solvent	Solubility	$\nu_{C=O}$ shift / cm^{-1}	SCXRD
Acetone	Soluble	No shift	Form I
Acetonitrile	Soluble	No shift	Form I
Chloroform	Insoluble	N/A	N/A
Dichloromethane	Insoluble	N/A	N/A
Ethanol	Soluble	No shift	Form I
Ethyl Acetate	Soluble	No shift	Form I
Hexane	Insoluble	N/A	N/A
Methanol	Soluble	No shift	Form I and II
Tetrahydrofuran	Soluble	No shift	Form I
Water	Insoluble	N/A	N/A

The polymorphic screen yielded one novel result; a new form of **58** derived from very slow evaporation from methanol. Form II crystallized concomitantly with Form I. In order to isolate Form II, the crystallization had to be performed in a narrow (12 mm) sample vial, which was capped, with a single pin-hole present. If these conditions were not met, for example, in large vials or uncapped small vials, only Form I was formed. The specific conditions under which Form II crystallizes resulted in difficulties scaling-up the crystallization, and so only small quantities of the material were prepared. Form II crystallizes in the space group

$P2_1/n$ [a = 16.397(4), b = 4.0573(8), c = 18.583(4), β = 102.101(5), V = 1208.8 Å^3]. This is the first true polymorph identified for **58** to-date. Atropisomerism around C_7-C_8 is not present in this structure, and only the major S(6) conformer is observed. Comparison of the theoretical PXRD patterns for both forms is shown in Figure 6.

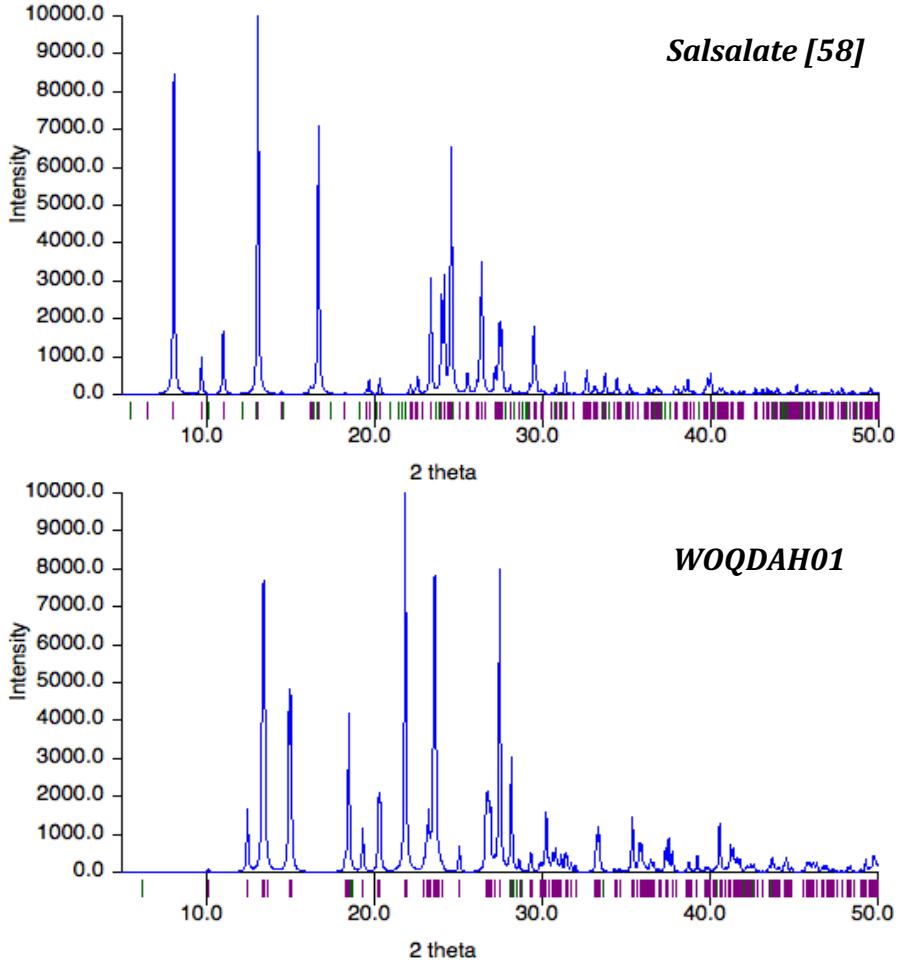


Figure 6: Theoretical comparison of PXRD patterns for 58 polymorph and WOQDAH01.

In a similar fashion to Form I, the main structural feature of Form II is the robust $R_2^2(8)$ hydrogen bonded dimer. Crystal growth extends along the *ac*-plane via this dimer interaction. In contrast to the very weak interactions that support growth in WOQDAH/WOQDAH01, this dimer interactions is coupled with a moderate (2.6 Å) C-H···O=C interaction between the ester carbonyl and an aromatic C-H (Figure 7).

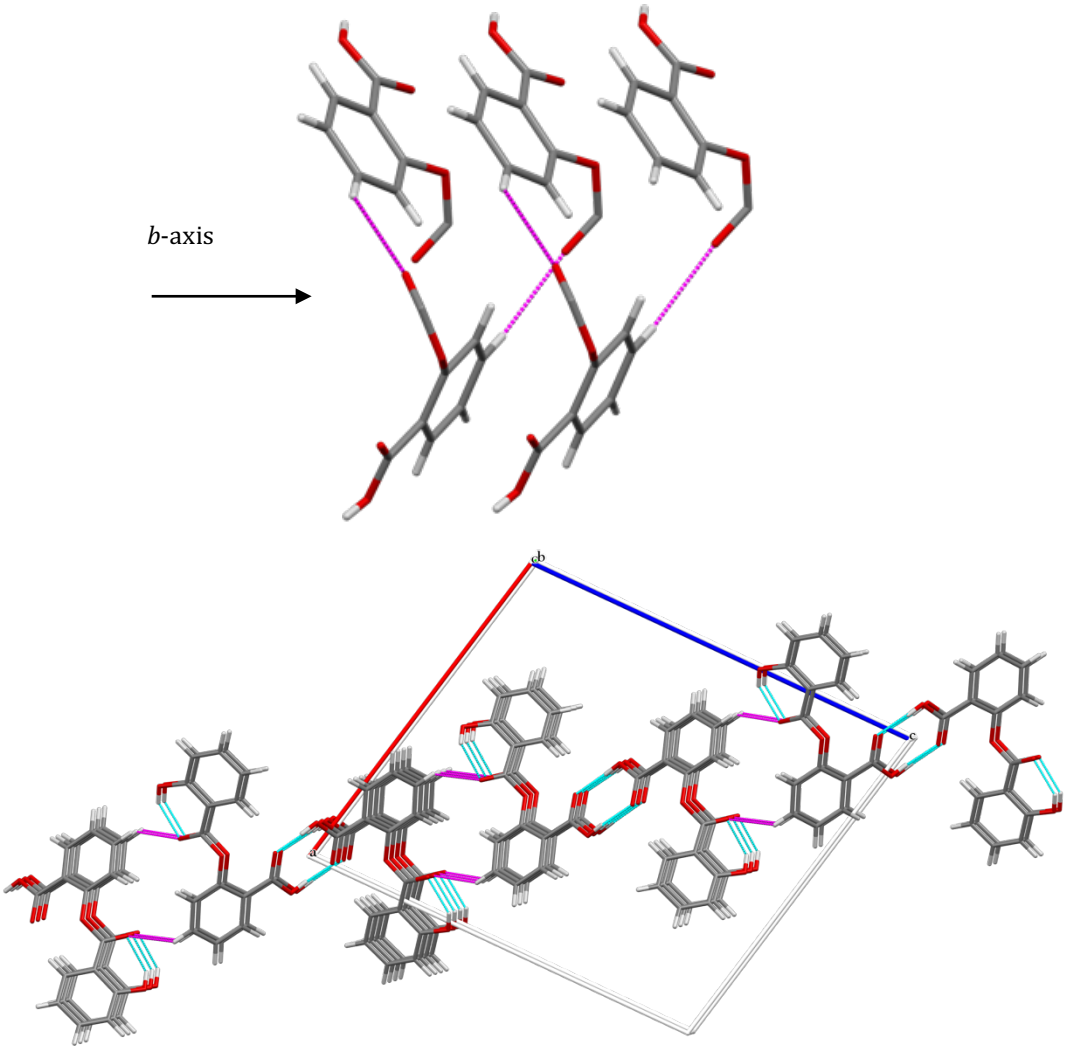


Figure 7: $R_2^2(8)$ dimer (cyan) and C-H···O=C (magenta) hydrogen bonds extending growth in the *ac*-plane of 58 polymorph [growth along *b*-axis in top left corner].

Comparison of the relative angles of the salsalate aromatic rings reveals further differences between the two forms. Form I presents with the rings closer to perpendicular with each other (78.1° in WOQDAH,[11] and 81.9° in WOQDAH01[3]), whereas Form II has the rings at just 42.1° to each other. This difference in angle allows the formation of the additional C-H···O=C interaction in Form II (Figure 7), contributing to a more linear progression of crystal growth than that observed in Form I.

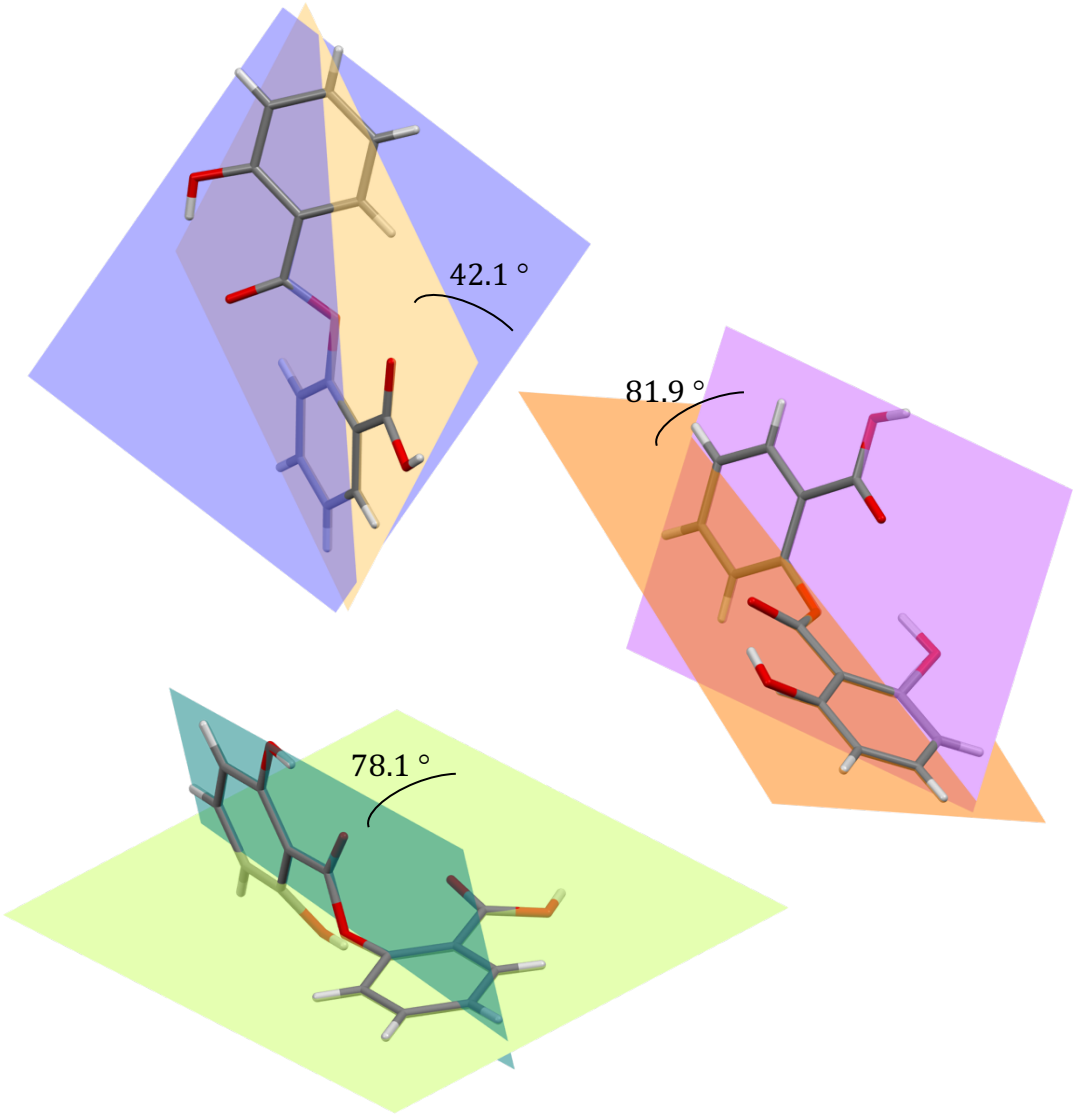


Figure 8: The relative angles of the aromatic rings in 58 Form II (top), WOQDAH (bottom) and WOQDAH01 (centre).

The already known form of **58** (Form I) has a distinctive habit, forming large block shaped crystals (Figure 9, left). The linear progression of crystal growth along the *ac*-plane in the new form (Form II), leads to a fine needle habit (Figure 9, right), that is easily distinguished from Form I.

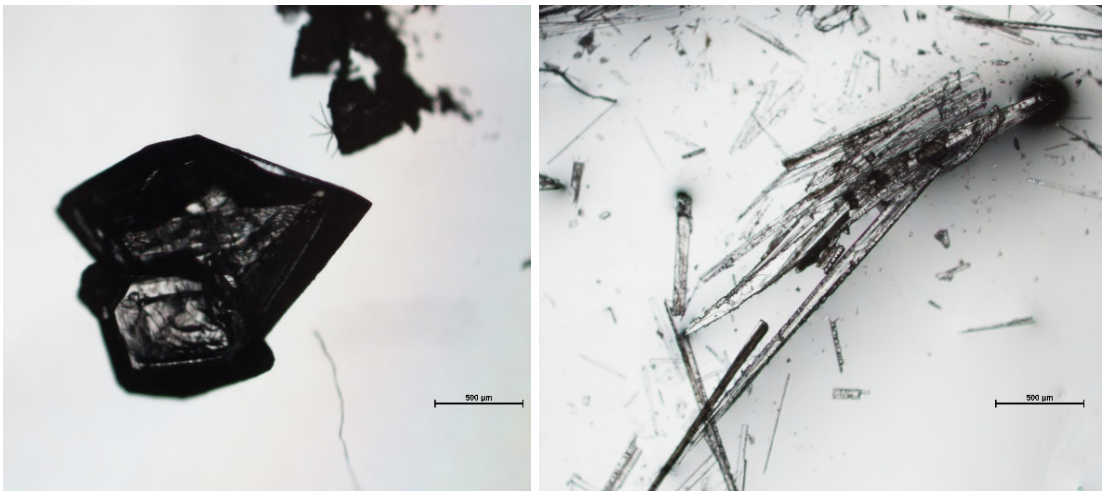


Figure 9: Comparison of the crystal habits of 58 [Form I left, Form II right].

The thermal behaviour of these materials was investigated using DSC and visual melting point analysis. The reported melting points for **58** range from 147 °C[13] to 225 °C[14] (although this report[14] stated that the melting point was 'mixed'), with most reports lying in the 140-150 °C range.[13,15] DSC analysis was performed for each of the concomitant forms from a methanol evaporation experiment. The needle habit (Form II) and the normal block shaped crystals (Form I) were carefully separated under a microscope before analysis.

Form I displays a broad melting point of approx. 130-140°. This is a similar range to the 131-145 °C determined in a study by Habgood *et al*. in 2013.[16] Form II displays a complex set of sharp endotherms in DSC at *ca*. 127 °C, 133 °C, 154 °C and 156 °C, all inconsistent with that observed in Form I, there is also a broad endotherm observed from 150-155 °C. In both cases, thermal decomposition of the material was observed to occur at temperatures greater than 200 °C. Due to the issues in scaling-up the crystallization as described above, preparation of sufficient sample quantities of Form II for analysis presented a challenge. DSC was performed on approx. 0.8 mg of material (typical DSC quantities are in the order of 5-7 mg), and so small thermal changes in the aluminium pan through the experimental sequence may have given rise to additional artefacts in the DSC pattern.

A visual melting point of **58** was also determined, utilising the needle shaped crystals obtained from methanol; melting was observed to occur at 147-149 °C.

It appears that the thermal behaviours of salsalate are complex, which would account for the various melting points reported in the literature. Nevertheless, the melting point recorded for **58** Form II is higher than that of Form I, indicating that the material is slightly more stable.

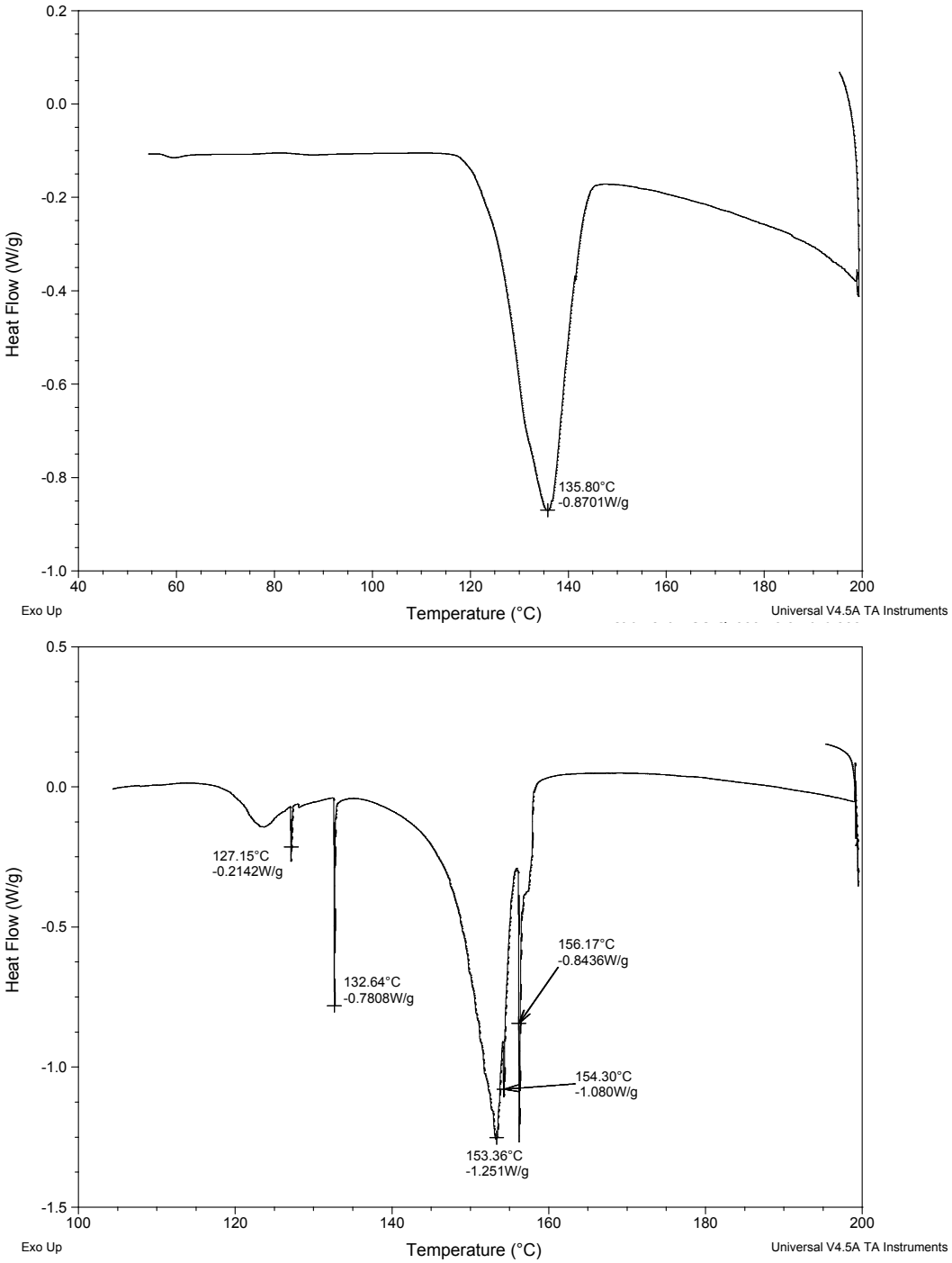


Figure 10: DSC analysis of 58 Form I (top) and Form II (bottom), performed at 2 °C/min.

3.3 Coformer Selection and Cocrystallization Screening

In consideration of **58** as an API, specific criteria were included in the knowledge-based coformer selection process.[17]

(i) Coformers would have to be GRAS (Generally Regarded As Safe).

(ii) Coformers would have to be accessible from fine chemical suppliers at a reasonable cost.

In terms of cocrystal design, efforts were directed toward breakage of the acid-acid homomeric dimer (Figure 3) via introduction of compatible functional groups (coformers). The ester and phenol functional groups were also considered as potential hydrogen bonding sites, keeping in mind the preferential formation of the S(6) intramolecular hydrogen bond, and atropisomerism around C_7-C_8 as shown previously (Figure 4).

The homomeric acid-acid or amide-amide dimer is less energetically favoured than the alternative acid-amide heterodimer.[18] This fact is commonly exploited during coformer selection for acid or amide materials.[18,19] There are 482 examples of the acid-amide $R_2^2(8)$ dimer in the CSD,[10] and this interaction would be predicted to form using Hunter's values.[20] Another functional group that was identified as a potential target was the pyridyl moiety, which could accept a hydrogen bond from either the phenol (1008 examples in the CSD),[10] or the carboxylic acid O-H (1549 examples in the CSD).[10] The COOH···N interaction has a high statistical occurrence, and high persistence in the presence of competing functional groups, with the alcohol O-H···N interaction slightly less reliable.[21] The final functional group that was selected as a target was the carboxylic acid of other molecules, with the aim of accessing the rarer acid-acid heterodimer bonding motif (66 examples in the CSD).[10]

Since there are no reported cocrystals of **58** to date, salicylic acid [**60**] was used in the CSD[10] searches as a model for **58**. There are 68 cocrystals structures in the CSD[10] containing **60** (where the search for salicylic acid is performed with requirement for the number of discrete molecules >1). The most common hydrogen bonding interaction observed for cocrystals of **60** are the acid-amide

$R_2^2(8)$ dimer (10 structures, 80% frequency), the acid-pyridyl $R_2^2(7)$ dimer (15 structures, 100% frequency), and the acid-pyridyl O-H···N-C hydrogen bond (23 structures, 92% frequency) [Figure 11].

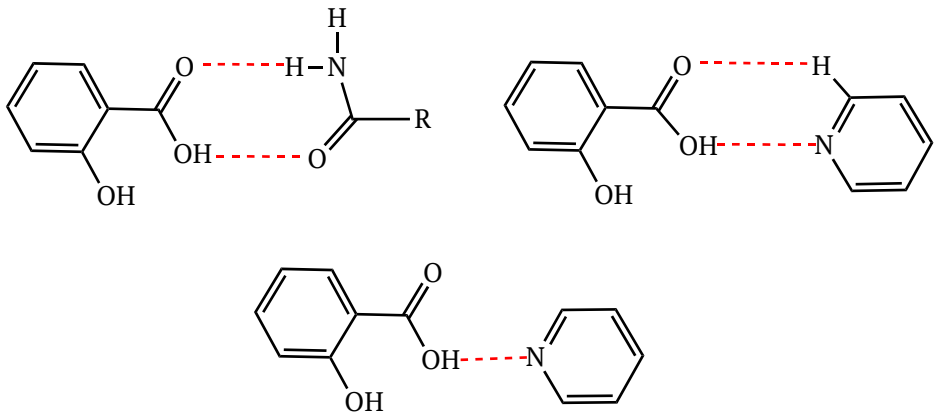


Figure 11: Interactions observed in cocrystals of 60.

Of the cocrystals investigated in the study of salicylic acid as a model, the polymorphism displayed in cocrystals of **60** with isonicotinamide [**61**] was of particular interest in the context of prediction of potential hydrogen bond interactions for **58**. There are two polymorphic structures of the **60.61** cocrystal published in the literature; in the first reported structure, XAQQEM [$P2_1/c$],[22] the full amide ladder from isonicotinamide is retained, and only capped at the ends by the acid-pyridyl interaction to the salicylic acid (Figure 12).

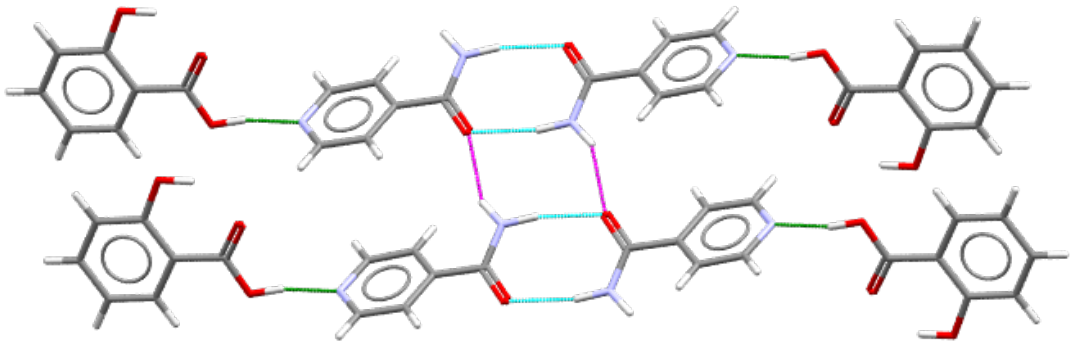


Figure 12: Persistence of the amide ladder in the 1:1 cocrystal 60.61 [magenta, cyan and green].[22]

The second structure, QAFTID [*P*-1],[23] determined in 2016 by Zhou *et al.*, has a 2:1 stoichiometric ratio of salicylic acid to isonicotinamide [**60.61**], and melts at a slightly higher temperature (135.6 °C versus 132.1 °C for XAQQEM)[22] suggesting that this form is the more stable of the two. The motifs displayed by this material are consistent with the suite of intermolecular interactions that

would be predicted using traditional methods, the $R_2^2(8)$ heteromeric acid-amide dimer and also the acid-pyridyl interaction (thus creating the 2:1 ratio) [Figure 13].

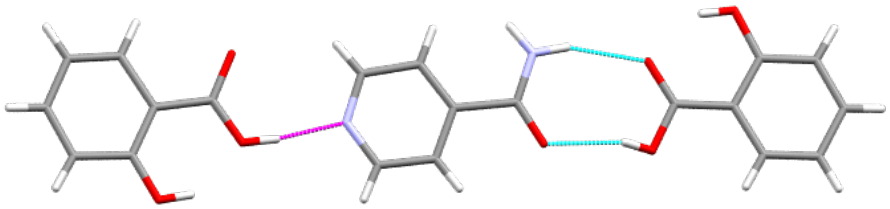


Figure 13: Interactions observed in 2:1 60.61 cocrystal QAFTID[23] [magenta and cyan].

This 3-molecule motif is almost planar, and each motif is connected to the next at an angle of 87.4° *via* bifurcation of the second salicylic carbonyl to the amide N-H [Figure 14].

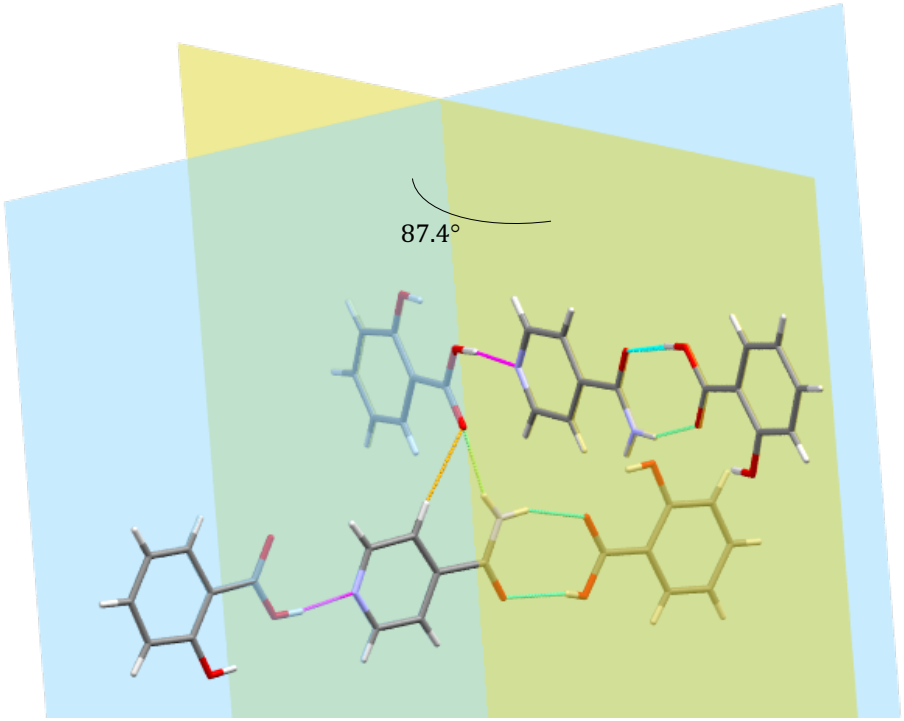


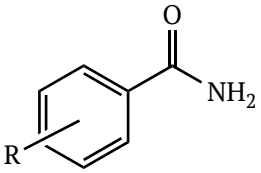
Figure 14: Planes of adjacent three molecule motifs observed in 2:1 cocrystal of 60:61 (QAFTID).[23]

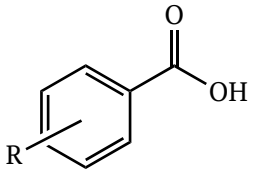
Considering the behaviour of salicylic acid [**60**] in these cocrystals, it was of particular interest to investigate how the additional steric bulk of salsalate could affect the hydrogen bonding motifs that may be observed in a potential cocrystal with isonicotinamide and/or other similar amides.

With all of these factors in mind, a concise library of 36 substituted aryl acids and amides was chosen for screening of **58**, encompassing the potential to form all four of the primary supramolecular synthons used in cocrystallization studies (Section 1.1.3, Figure 9). The limited number of pyridine-containing materials contained on the GRAS list meant that only four such compounds were included for screening (Table 3).

Screening was performed using traditional solid state grinding, followed by analysis of the powdered products *via* PXRD and IR. Of the 32 materials used for screening, all but one (isonicotinamide, **61**) presented negative results for cocrystallization, with no $\nu_{c=o}$ shift observed in the IR spectrum, and the PXRD pattern of the grinding product simply a combination of the two coformers.

Table 3: Library of coformers used in this study (successful coformer shown in green).





#	Amide	#	Acid
13	Benzamide	**76**	Benzoic acid
61	Isonicotinamide	**77**	2-Aminobenzoic acid
1	Nicotinamide	**78**	3-Aminobenzoic acid
62	2-Aminobenzamide	**2**	4-Aminobenzoic acid
63	3-Aminobenzamide	**79**	2-Methylbenzoic acid
64	4-Aminobenzamide	**80**	3-Methylbenzoic acid
9	2-Methylbenzamide	**81**	4-Methylbenzoic acid
65	3-Methylbenzamide	**60**	2-Hydroxybenzoic acid
66	4-Methylbenzamide	**82**	3-Hydroxybenzoic acid
67	2-Hydroxybenzamide	**83**	4-Hydroxybenzoic acid
68	3-Hydroxybenzamide	**84**	2-Fluorobenzoic acid
69	4-Hydroxybenzamide	**85**	3-Fluorobenzoic acid
70	2-Fluorobenzamide	**86**	4-Fluorobenzoic acid
71	3-Fluorobenzamide	**87**	2-Amidobenzoic acid
72	4-Fluorobenzamide	**88**	3-Amidobenzoic acid*
73	2-Amidobenzamide	**89**	4-Amidobenzoic acid*
74	3-Amidobenzamide*	**90**	Isonicotinic acid
75	4-Amidobenzamide*	**91**	Nicotinic acid

* Not available at a reasonable cost and so were excluded;

The IR spectrum of the grinding product of **58** and isonicotinamide [**61**] showed considerable shifts for both the carbonyl C=O (increase of -14 cm^{-1} for the amide, +4 cm^{-1} for the carboxylic acid), and the amide N-H (increase of 10 cm^{-1} and 26 cm^{-1} respectively) [Figure 16 and Figure 17]. There was also appearance of a new PXRD pattern, inconsistent with either of the reagents (or any known polymorphs thereof), indicating formation of a new solid form (Figure 15) These results were interpreted as strong indication for the formation of a 1:1 cocrystal.

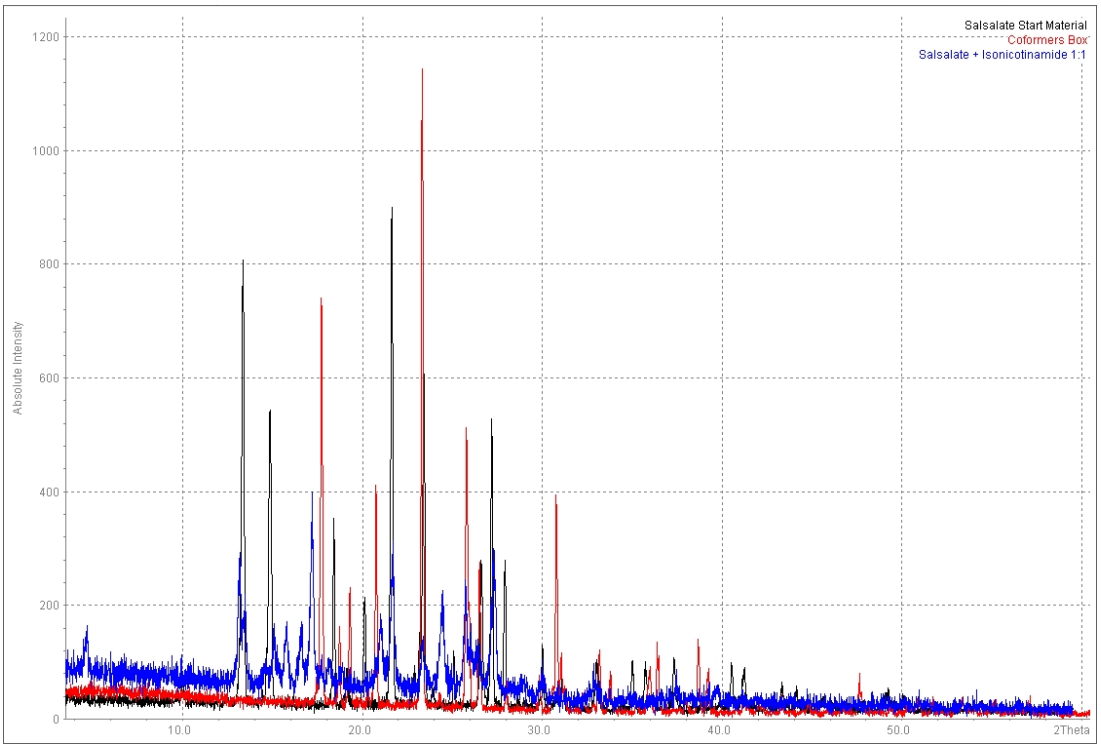


Figure 15: The PXRD comparison of 58 [black], 61 [red] and the ball mill grinding product [blue].

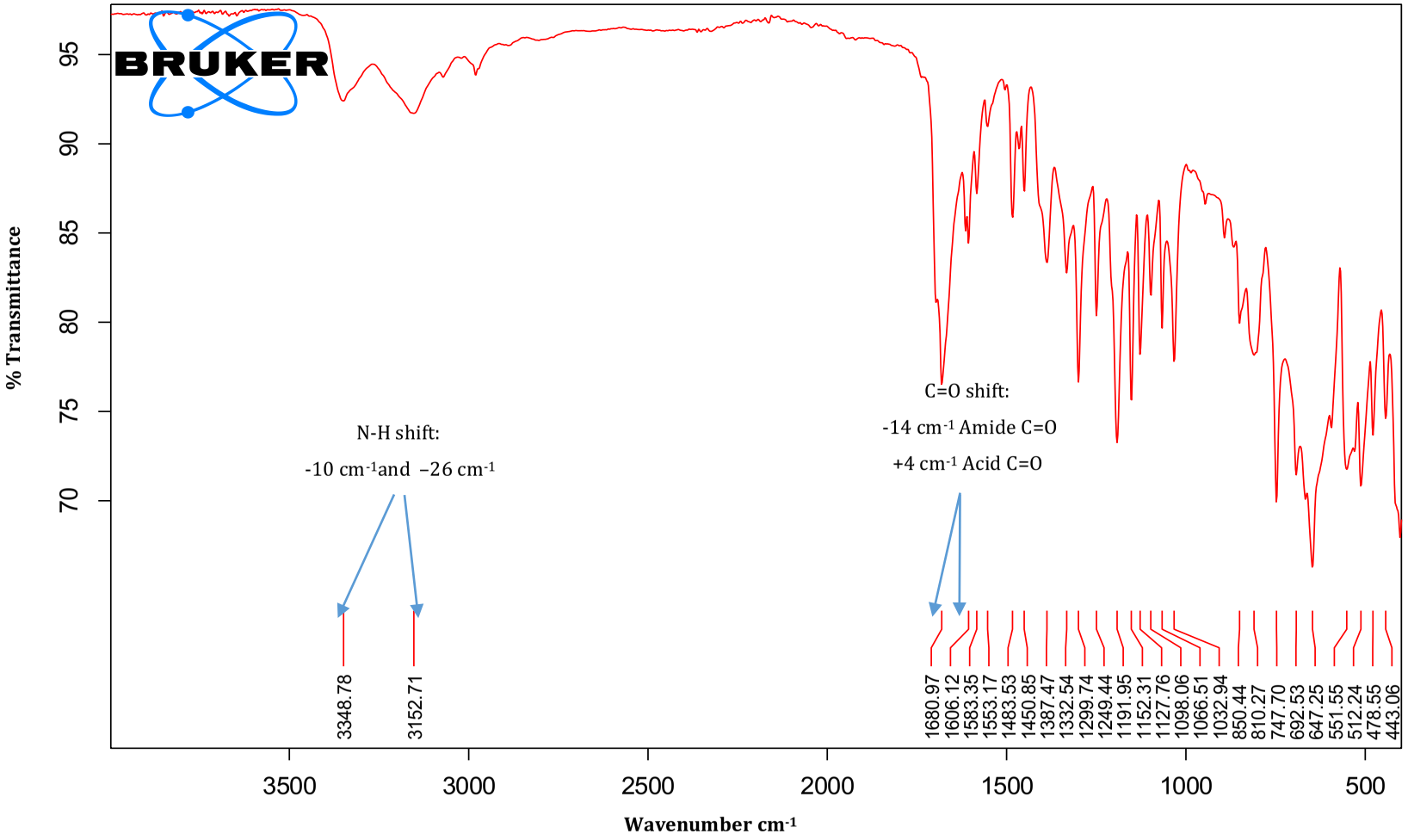


Figure 16: Infrared spectrum of salsalate- isonicotinamide cocrystal [58.61].

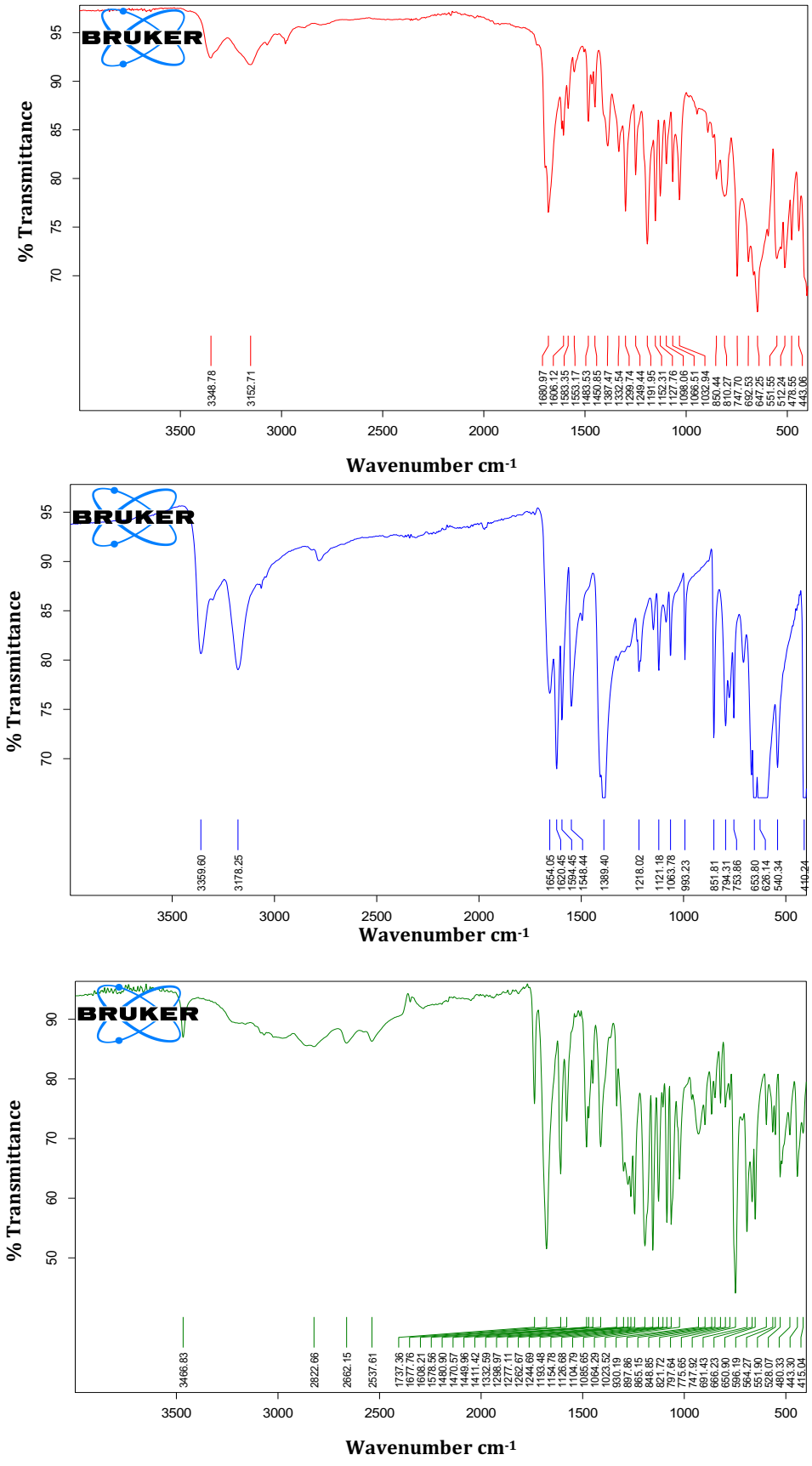


Figure 17: Comparison of IR patterns for salsalate-isonicotinamide cocrystal [58.61, red], isonicotinamide [61, blue] and salsalate [58, green].

Single crystal analysis of the **58.61** cocrystal was performed on crystals obtained by slow evaporation from acetonitrile, which crystallizes in the space group *P*-1 [a = 6.9660(12), b = 7.1847(14), c = 18.983(4), α = 94.277(5), β = 95.816(4), γ = 104.860(4), V = 908.6(3) Å^3]. The cocrystal is characterised by two primary hydrogen bonding features. The first feature is the $R_2^2(8)$ homomeric-dimer retained from the parent isonicotinamide crystal structure [Figure 18], and this is capped at the sides by acid-pyridyl hydrogen bonding from the salsalate carboxylic acid to the pyridine nitrogen of the isonicotinamide. The isonicotinamide dimer motif observed here is consistent with that observed in similar cocrystals of isonicotinamide and carboxylic acids, as reported previously by Aakeröy.[24]

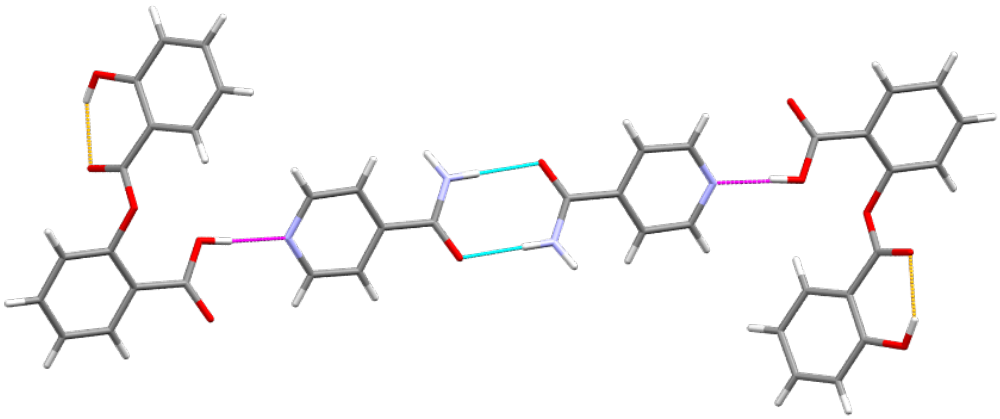


Figure 18: Primary hydrogen bonding interactions in 58.61 cocrystal [magenta, cyan and orange].

The secondary feature of the bonding in the salsalate cocrystal [**58.61**] is a discrete motif formed by hydrogen bonding from the second amide N-H to the carboxylic acid carbonyl oxygen, and the intermolecular S(6) ring motif observed in **58** (Figure 19).

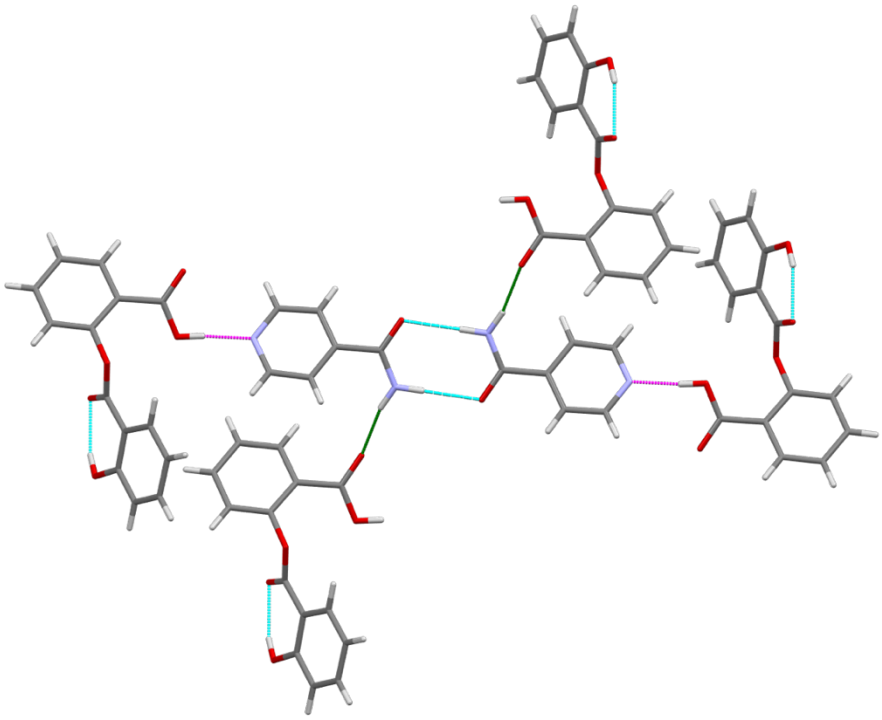


Figure 19: Hydrogen bonding interactions observed in 1:1 cocrystal of 58.61 [magenta, cyan and green].

The structural motifs observed in the salsalate:isonicotinamide cocrystal [**58.61**] reflect, in part, both of the known cocrystals of salicylic acid [**60**] with isonicotinamide [**61**], discussed above. The **61** homodimer is retained, similar to that observed in XAQQEM,[22] but the amide ladder is broken, and the second amide N-H forms hydrogen bonds to the salicylic acid instead, similar to QAFTID.[23] It is interesting to see that in all of these structures the acid-pyridyl interaction prevails, suggesting that this is the most strongly dependable motif that can be exploited for cocrystallization with salsalate.

DSC analysis of the cocrystal [**58.61**] displays a strong endotherm at 126-128 °C, which is lower than each of the reagents (155-157 °C for **61**, and 147-149 °C for **58**)[13,15,25] [Figure 20]. A study in 2015 showed that cocrystals most commonly form with a melting point between that of the individual coformers (55.3%), while a melting point lower than both coformers is seen less frequently (15.8%).[26] Decomposition is observed to begin at approximately 190 °C, which is consistent with decomposition of **58**.[27]

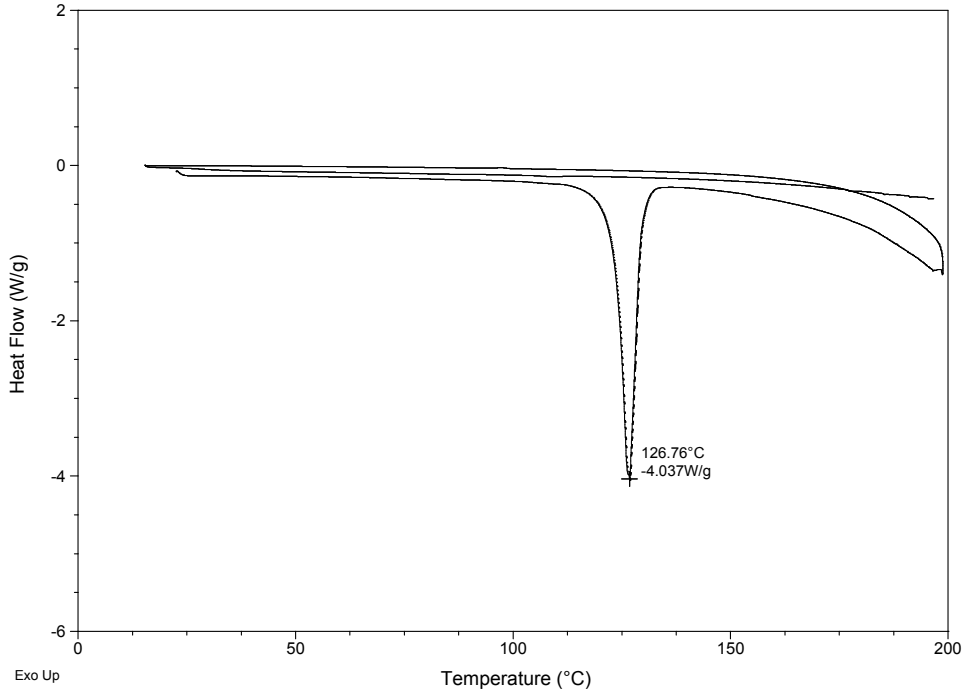


Figure 20: DSC endotherm for 58.61 cocrystal.

The cocrystallization success rate for **58** was unexpectedly low based on the functional groups present. The literature would suggest that acid-amide interactions would occur here, likely in preference over the acid-pyridyl interaction, and although one can never guarantee the formation of a cocrystal, the examples in the literature created anticipation of a far higher level of success with the set of coformers designed for this study.

It was decided to take the project in two directions going forward:

(i) To build upon the successful result gained in initial screening and further investigate the landscape of the acid-pyridyl interaction for **58**.

(ii) To examine the low success rate of the initial 32-compound screen by analysing the cocrystallization behaviours of the products obtained from neat grinding experiments from solution crystallization.

3.4 Investigation of the Acid-Pyridyl Interaction

Investigation of the acid pyridyl interaction required a separate library of pyridyl and pyrimidyl coformers (Figure 21). This secondary screen necessitated the removal of the requirement for GRAS coformers in order to create a large enough library for screening. 13 materials were selected for screening based on availability: acridine [**92**], 2-amino-4-chloro-6-methylpyrimidine [**93**], 4-amino-4,6-dimethylaminopyrimidine [**94**], 2-aminopyrimidine [**95**], 4-aminopyridine [**96**], 4,4'-bipyridine [**15**], 4-hydroxypyridine [**97**], 2-pyridinecarbonitrile [**98**], 3-pyridinecarbonitrile [**99**], 4-pyridine-carbonitrile [**100**], 2,6-pyridinecarboxamide [**101**], 2-pyridinethioamide [**102**], and 4-pyridinethioamide [**103**].

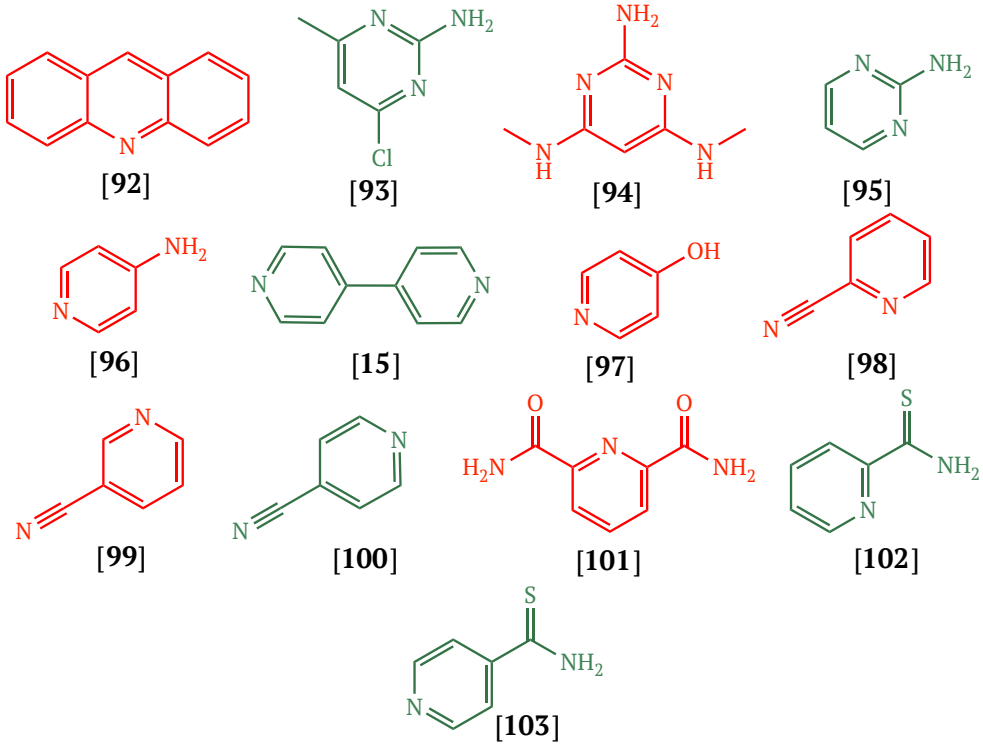


Figure 21: The library of pyridyl and pyrimidyl coformers used for screening (materials displaying evidence of structural change shown in green).

The investigation was carried out in the same way as has been previously described (Section 3.3). Of the 13 materials screened in this secondary study, 6 grinding products showed evidence of the formation of a new solid form: **93**, **95**, **15**, **100**, **102** and **103**. Efforts were directed to characterising the solid state bonding motifs of these materials using SCXRD. Each material was dissolved in ethanol, isopropanol, acetonitrile, and ethyl acetate. Methanol was not included

as a solvent in this screen for reasons that will be discussed later in Section 3.6.1. The solvent was allowed to evaporate over approximately 1 week, yielding crystals suitable for analysis in 3 cases. The remaining 3 materials were confirmed as novel solid forms using DSC.

The first material successfully characterised using SCXRD was a cocrystal of **58** with 4,4'-bipyridyl [**15**]. **15** was chosen as a simple model to investigate the acid-pyridyl interaction without the added competition of other hydrogen bond donors/acceptors. **15** was also selected because it is a commonly used coformer in the cocrystallization of acid materials (223 cocrystals in the CSD[10]). It was postulated that the symmetry of **15** could mimic the symmetrical position of the $R_2^2(8)$ dimer at the inversion centre of the previous crystal structure (**58.61**, Figure 22).

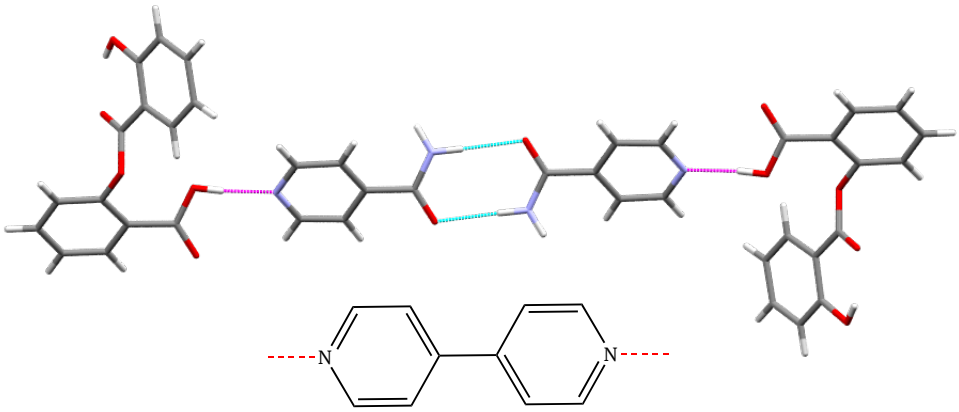


Figure 22: Comparison of symmetrical relationship of 4,4'-bipyridyl [15] with the isonicotinamide dimer [61].

Crystal structure analysis of crystals obtained from ethanol showed this hypothesis was correct, [*P*-1, a = 7.6250(12), b = 7.9985(11), c = 13.368(2), α = 95.770(5), β = 97.757(5), γ = 94.778(5), V = 799.9(2) Å^3]. Thus, the ratio of materials in this cocrystal is 2:1, with the two pyridyl moieties capped at either end by **58** molecules. The intermolecular $S(6)$ hydrogen bond is retained, but an additional C-H···π interaction is observed between the molecules of **58** (Figure 23).

There is a weak C-H···C=O hydrogen bond between the carbonyl oxygen atom of the carboxylic acid and the hydrogen *para* to the ester on the phenol ring. This creates zig-zag sequences of molecules in the *bc*-plane, and the molecules of **15**

lie in channels between the **58** molecules (Figure 24). The aromatic rings in **58** are held almost perpendicular to each other (relative angle of the rings 88.7°) [Figure 25].

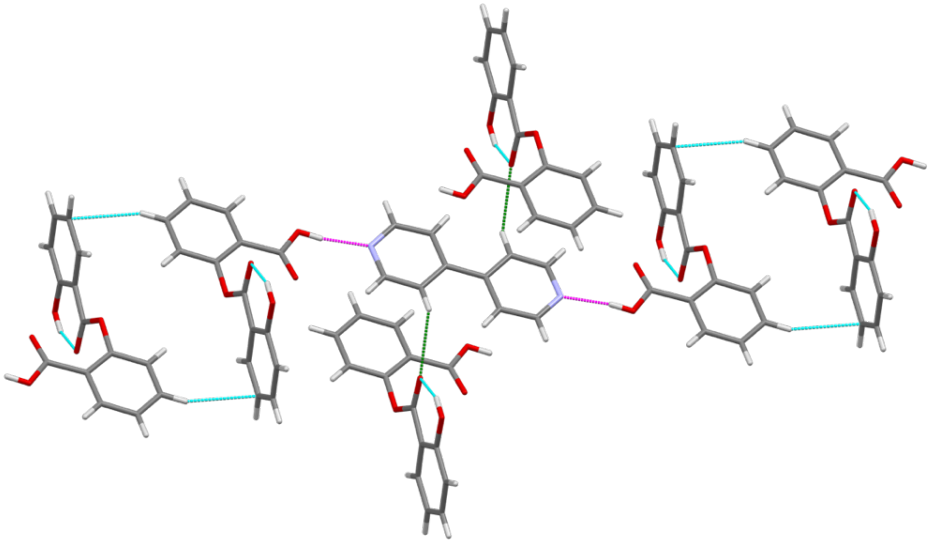


Figure 23: Hydrogen bonding interactions observed in 58.15 2:1 cocrystal [bonds in magenta, cyan and green].

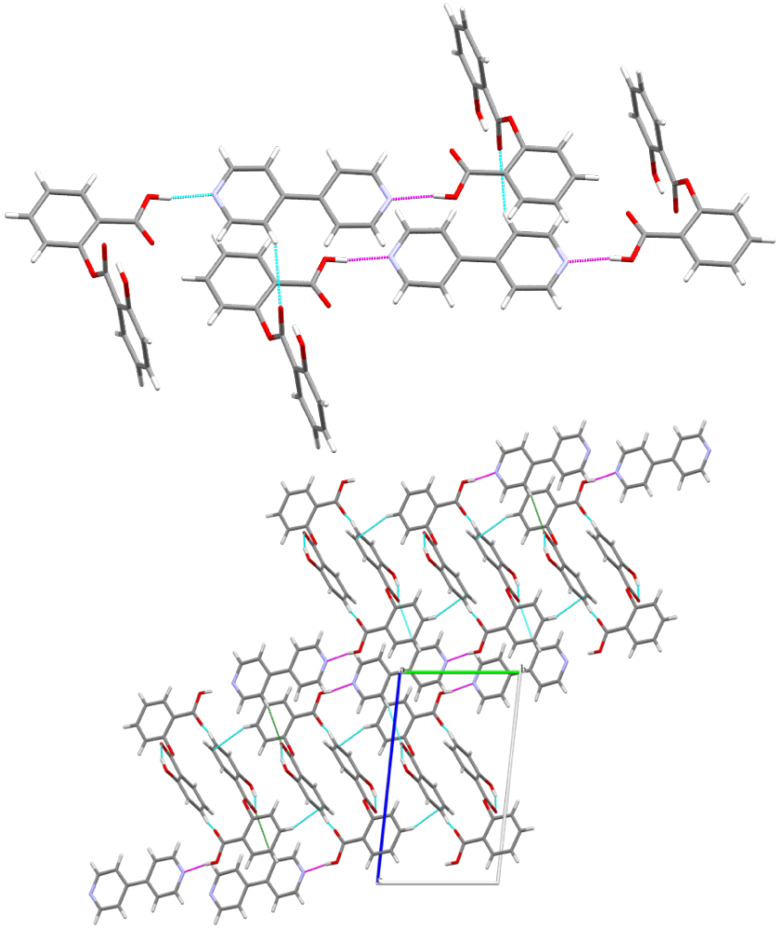


Figure 24: Zig-zag of molecules in *bc*-plane of 2:1 cocrystal 58.15 [bonds in magenta, cyan and green].

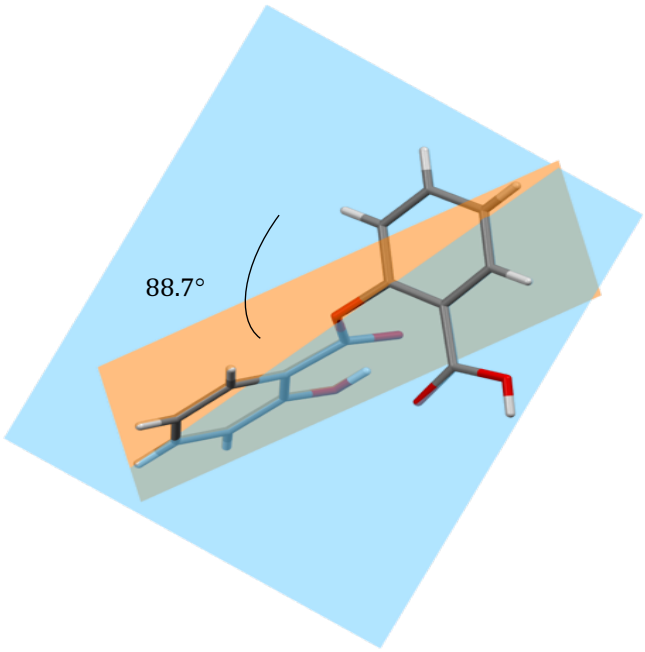


Figure 25: Relative angle of aromatic rings in 58.51 cocrystal.

The second material successfully characterised in this series was the 1:1 cocrystal of **58** with 2-amino-4-chloro-6-methylpyrimidine [**93**], *P*-1 [a = 7.6518(3), b = 9.4553(4), c = 14.2495(6), α = 83.969(2), β = 74.507(2), γ = 72.482(2), V = 947.07(7) Å^3]. The bonding motifs observed in this material are more interesting, due to the greater variety of hydrogen bond donors and acceptors in the coformer. The primary interaction observed in the material is an $R_2^2(8)$ dimer formed between the carboxylic acid of **58**, one of the amine N-H atoms and one of the two pyrimidine nitrogens (Figure 26). The other pyrimidine nitrogen and amine N-H atom participate in a further $R_2^2(8)$ homomeric dimer interaction to a second molecule of **93,** which in turn hydrogen bonds back to a second molecule of **58**, creating a heteromeric tetramer at a binary level. The pyrimidine homodimer lies in a symmetrical position between the two salsalate molecules, similar to that observed previously in the literature examples for isonicotinamide, and in the 1:1 cocrystal of **58.61**.

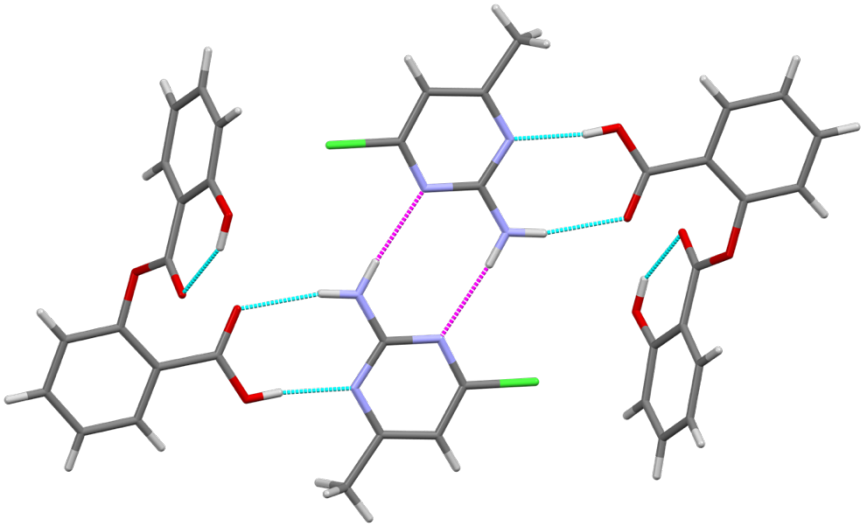


Figure 26: Symmetrical $R_2^2(8)$ dimers formed in 58.93 cocrystal (homomeric dimer in magenta, heteoromeric dimer in cyan).

Interestingly, the binary level core at the centre of the tetramer lies completely planar. All of the rings involved in the dimers lie perfectly flat in the same plane, with the benzoic acid rings just slightly twisted at just 9°; the phenol rings as mentioned previously, lie perpendicular above and below this plane (Figure 27).

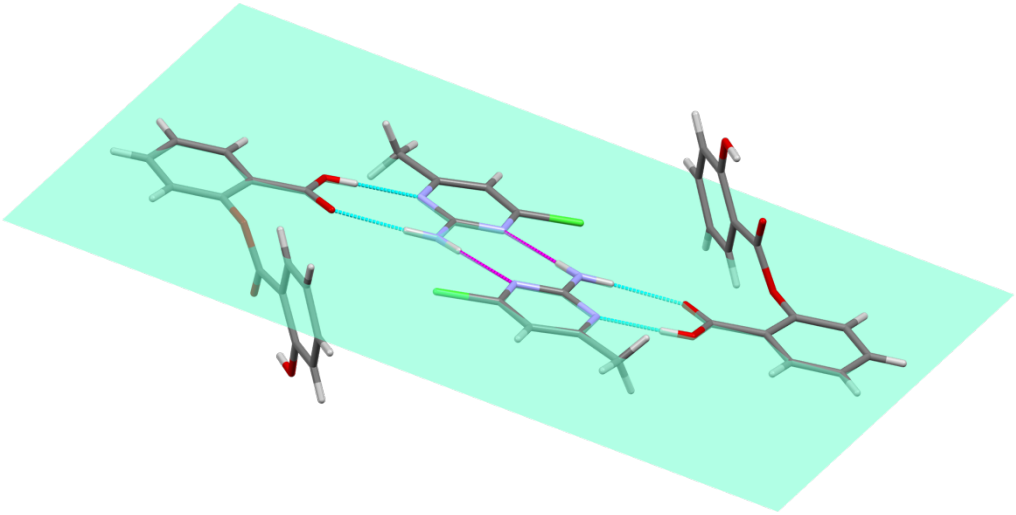


Figure 27: Planarity of the dimer interactions oberved in 58.93 cocrystal.

The tetrameric structure observed in this cocrystal structure reflect those observed in the study published by Ebenezer *et al*. in 2011.[28] The study determined the cocrystalline structures of 10 acids with **93**, all of which presented a high level of isostructurality through common bonding motifs.

The formation of the planar tetramer comprised of two forms of $R_2^2(8)$ homomeric and heteromeric dimers was persistent in all structures determined

by Ebenezer, and provided a robust supramolecular platform on which to design molecular architectures such as **58.93**. There are two structures determined in the above study that crystallize in the same *P*-1 space group, EXEYUC[28] (4-nitrobenzoic acid, **104**, Figure 28), and EXIBET[28] (4-methoxybenzoic acid, **105**, Figure 29). The structures are observed to be remarkably similar when visually examined in Mercury.[29]

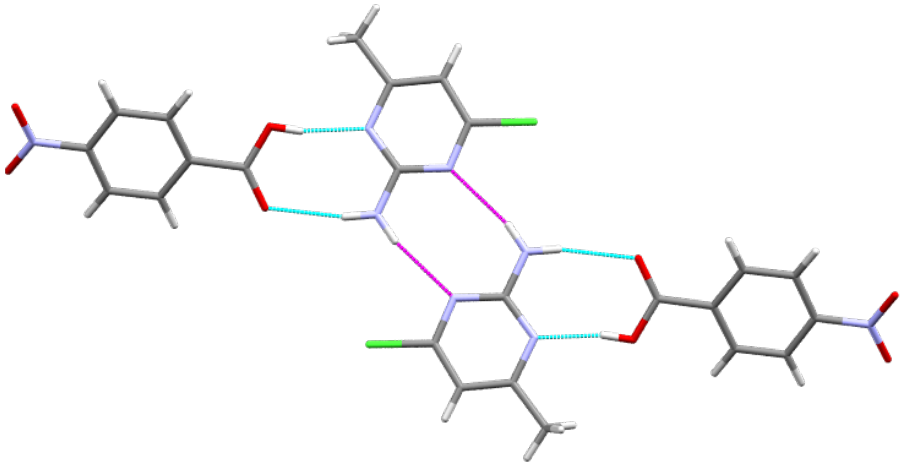


Figure 28: Persistence of structural motifs in EXEYUC [93.104].[28]

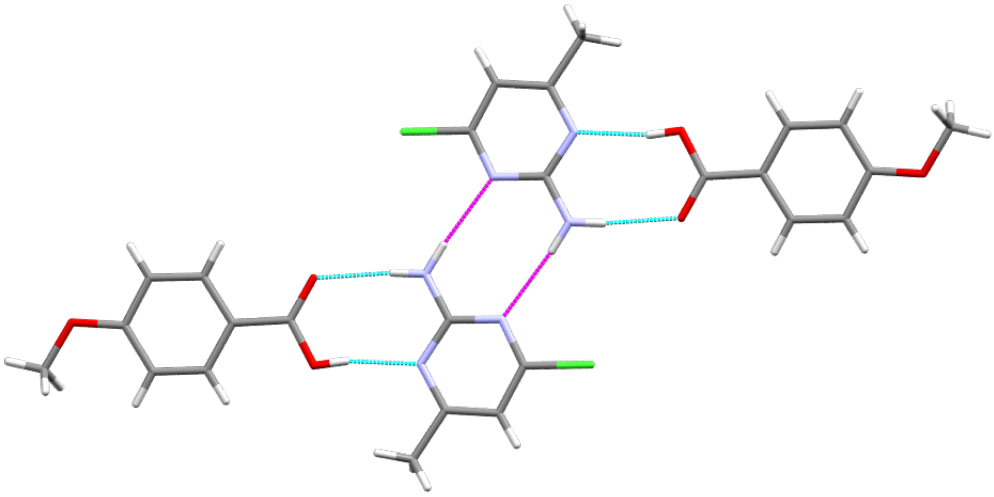


Figure 29: Persistence of structural motifs in EXIBET [93.105].[28]

Crystal growth in the **58.93** cocrystal is supported by a weak C-H···O-H interaction between these two phenol rings on the **58** molecules. There is much debate in the literature as to the validity of C-H···O-H interactions, as has been discussed previously within this thesis (Section 2.4). In this case, it appears that the C-H···O-H interactions observed here are structure defining because it supports formation of a ladder-like motif at a supramolecular level, extending growth in a linear fashion along the *a*-axis, interlinked by the planar dimer

interactions (Figure 30). This ladder motif is comprised of $R_4^4(26)$ tetramers (Figure 31).

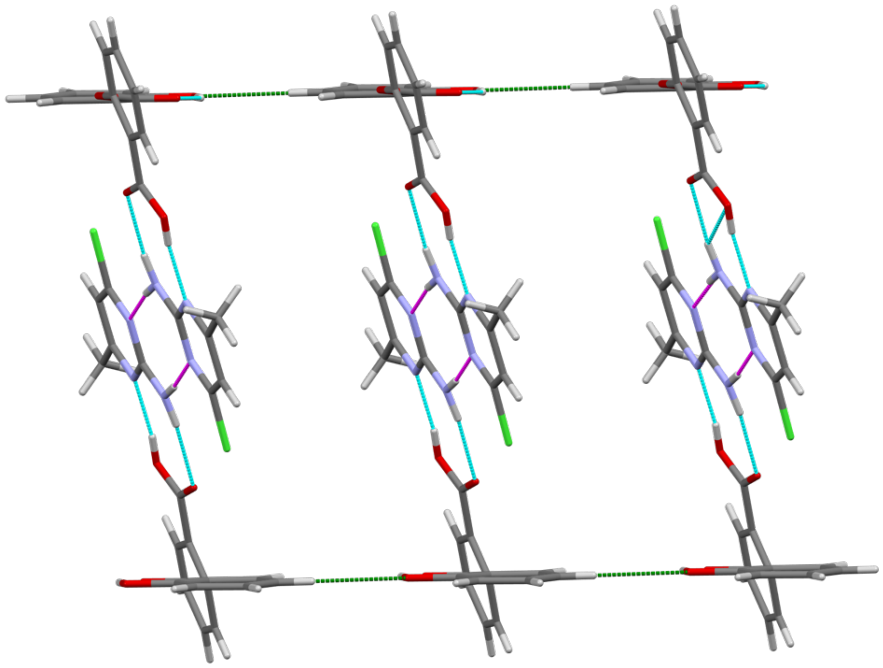


Figure 30: Ladder type formation at supramolecular level in 58.93 cocrystal (hydrogen bonds in magenta, cyan, and green).

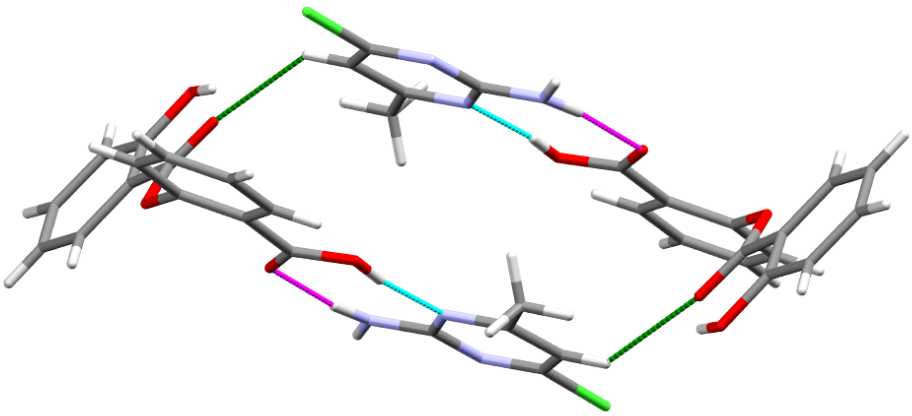


Figure 31: Formation of complex $R_4^4(26)$ tetramers in 58.93 cocrystal (hydrogen bonds in magenta, cyan, and green).

The final cocrystal that was successfully characterised in this series was that of **58** with 4-pyridinecarbonitrile [**100**] in a 1:1 ratio, [*Pc*, a = 15.7904(14), b = 7.2513(6), c = 31.105(3), β = 100.760(2), V = 3498.9(5) Å^3]. Crystallization of this material was challenging, the crystallizations from ethanol, acetonitrile and ethyl acetate did not successfully yield any crystals suitable for analysis. Only crystallization experiments from isopropanol were successful, although more

than one attempt at crystallization was required. The crystals formed very distinct plates, which were consistently twinned and therefore, presented a significant challenge in structure determination. Despite multiple attempts at structure solution, the best structure that could be determined has an R-factor of just 10.9%.

This material presents with a slightly different set of bonding motifs than those observed for the previous materials, with two different interactions observed between salsalate and the pyridyl moiety here. The first hydrogen bond observed is the targeted acid-pyridyl interaction, but there is also an acid-pyridyl $R_2^2(7)$ dimer hydrogen bonded motif (Figure 32). These motifs combine with two C-H···O=C interactions at a binary level to form a $R_4^4(16)$ tetramer. The salsalate aromatic rings adopt the perpendicular relationship that has been observed previously here also, with the two rings held at approximately 89° relative to each other.

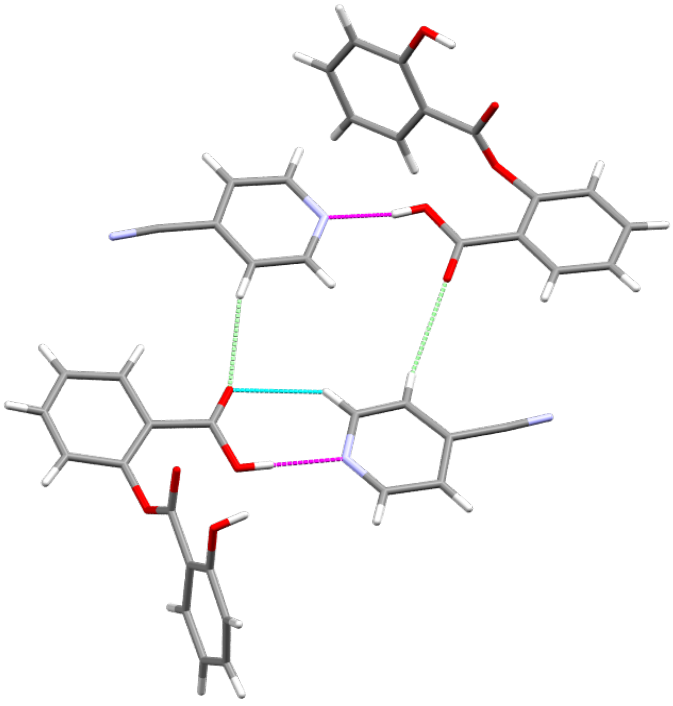


Figure 32: Acid-pyridyl $R_2^2(7)$ dimer in the structure of 58.100.

The nitrile moiety accepts a hydrogen bond from the nearest benzoic C-H of salsalate, forming a chain down the *c*-axis. Perpendicularly, the phenolic O-H, which sits in the major S(6) conformation, accepts a hydrogen bond from an

aromatic C-H on the next phenol ring, forming a C(6) chain down the *a*-axis (Figure 33).

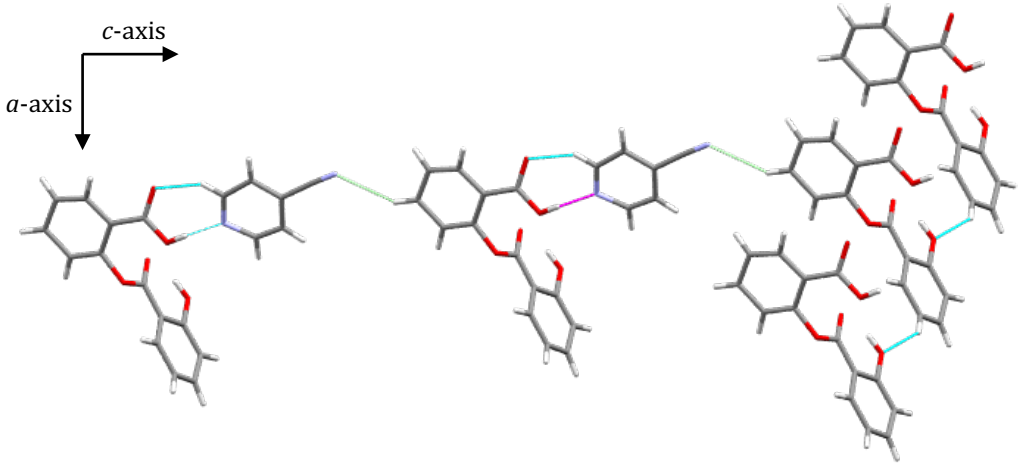


Figure 33: *C*(6) chains of C-H···N hydrogen bonding observed in 58.100 cocrystal.

The combination of the two C(6) chains in this cocrystal results in almost perfectly square, plate shaped crystals, which were quite twinned and presented a significant challenge in terms of structure determination from SCXRD.

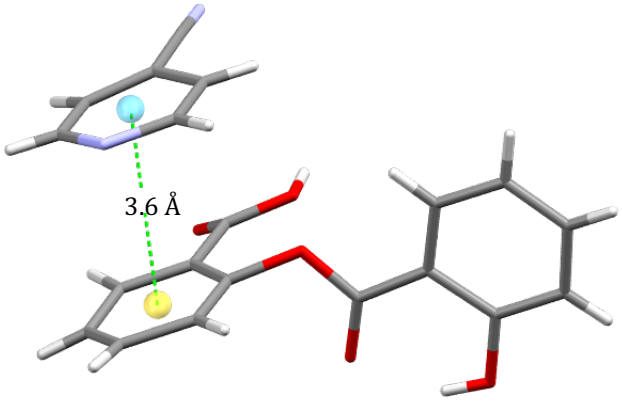


Figure 34: Evidence of π···π stacking interactions in 58.100 cocrystal.

Interestingly, this is the only structure in this series where π···π stacking is observed. Each benzoic acid moiety on salsalate is stacked to a pyridine ring of **100** at a distance of 3.6 Å and offset by approx. 22° (Figure 34). The pyrimidine cocrystals characterised by Ebenezer[28] showed a consistent π···π stacking interaction between the pyrimidine moieties. However, there does not appear to be any other evidence of such interactions in the other materials that were characterised in this study.

3.4.1 Thermal Analysis of Salsalate-Pyridine Cocrystals

DSC analysis of all 6 novel solid forms was undertaken. PXRD analysis of the neat grinding product of **58.95** had indicated a 2:1 solid state stoichiometric ratio of salsalate:2-aminopyrimidine. The grinding experiment was repeated with the correct ratio, and this material was crystallised and used for DSC analysis as the crystals obtained were not suitable for SCXRD. In each case, the DSC endotherm was sharp and distinctly different than those of the individual coformers (Table 4).

Table 4: DSC endotherms observed for salsalate cocrystals (DSC).

Compound Number	DSC Endotherm	m.p. of Coformer
2-Amino-4-chloro-6-methylpyrimidine [**93**]	137-141 °C	182-183 °C[30]
2-Aminopyrimidine (2:1)[a] [**95**]	105-107 °C	126 °C[31]
4,4'-Bipyridyl [**15**]	61-66 °C	112 °C[32]
4-Pyridine carbonitrile [**100**]	102-104 °C	78-80 °C[33]
2-Pyridinethioamide [**102**]	110-113 °C	137 °C[34]
4-Pyridinethioamide [**103**]	123-127 °C	210 °C[35]

a - DSC analysis of neat grinding product was undertaken as SCXRD analysis showed only coformers in the solution crystallization experiments;

4 of the 6 cocrystals display a melting point lower than both of the reagents, as observed for the salsalate:isonicotinamide cocrystal, discussed above [**58.61**]. For the remaining two cocrystals, **58.15** and **58.100** display melting points between those of the individual reagents, which is more commonly observed in the formation of cocrystals.[26] The results of the thermal analysis, coupled with the absence of N-H stretches in IR (which would indicate the formation of a salt) have confirmed the presence of the six new novel cocrystals of salsalate. Further efforts into alternative crystallization techniques would be required to determine the crystal structures of the remaining three materials.

3.5 Solution Crystallization Screening

As discussed above, the initial cocrystallization screen yielded a single cocrystal from the 32 solid state grinding experiments. In order to confirm these results, a full screen of the 32 coformers was performed using solution crystallization from each of methanol, ethanol, and acetonitrile. These solvents were chosen because (i) the coformers are soluble in these systems, and (ii) they are not overly volatile which should allow for growth of good quality single crystals, and (iii) they represent both protic and aprotic systems.

The products from section 3.3 were each dissolved in one solvent in an 18 mm glass sample vial and were allowed to stand until all of the solvent had evaporated (1-2 weeks). Each sample that yielded suitable crystals was then analysed via single crystal analysis, and the unit cells compared with those of the coformers. This secondary screen did not identify any new cocrystals of **58** with any of the 32 coformers.

Interestingly, for two of the materials investigated in this secondary screen, hydrolysis of the salsalate [**58**] was observed and a novel cocrystal of the resultant salicylic acid [**60**] with the coformer was characterised, in contrast to the grinding experiments. Two other coformers (4-hydroxybenzamide, **69** and nicotinic acid, **91**), showed evidence of the formation of salicylic acid in solution, but no cocrystals were identified in the vials.

The first of the two cocrystals was a 1:1 cocrystal of salicylic acid [**60**] with 4-methylbenzamide [**66**]. This cocrystal was determined in the *Pbca* space group [a = 8.2953(8), b = 9.9512(9), c = 33.538(4), V = 2768.5 Å^3]. The primary hydrogen bonding motif observed in this material is the acid-amide $R_2^2(8)$ heterodimer, one of the main target interactions of the original study. There is also an amide C(4) chain interaction down the *a*-axis, interlinking and creating an almost orthogonal relationship between the (almost) planar dimers (Figure 35 and Figure 36). C-H···O-H interactions from the *para*-tolyl group to the phenol (at approximately 2.4 Å), also stabilise the structure.

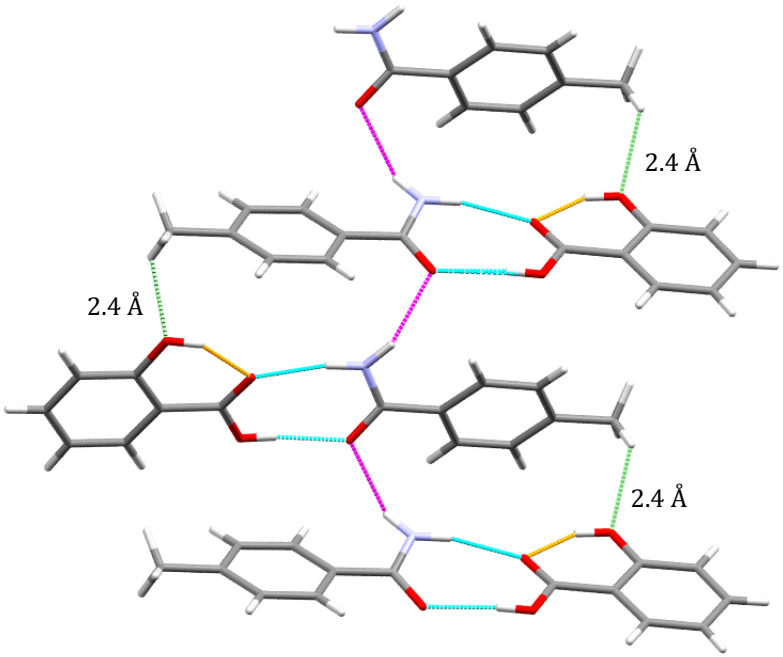


Figure 35: Acid-amide $R_2^2(8)$ heterodimer (cyan)and amide C(4) chain in 60.66 cocrystal (magenta), and additional C-H···O-H interactions (green).

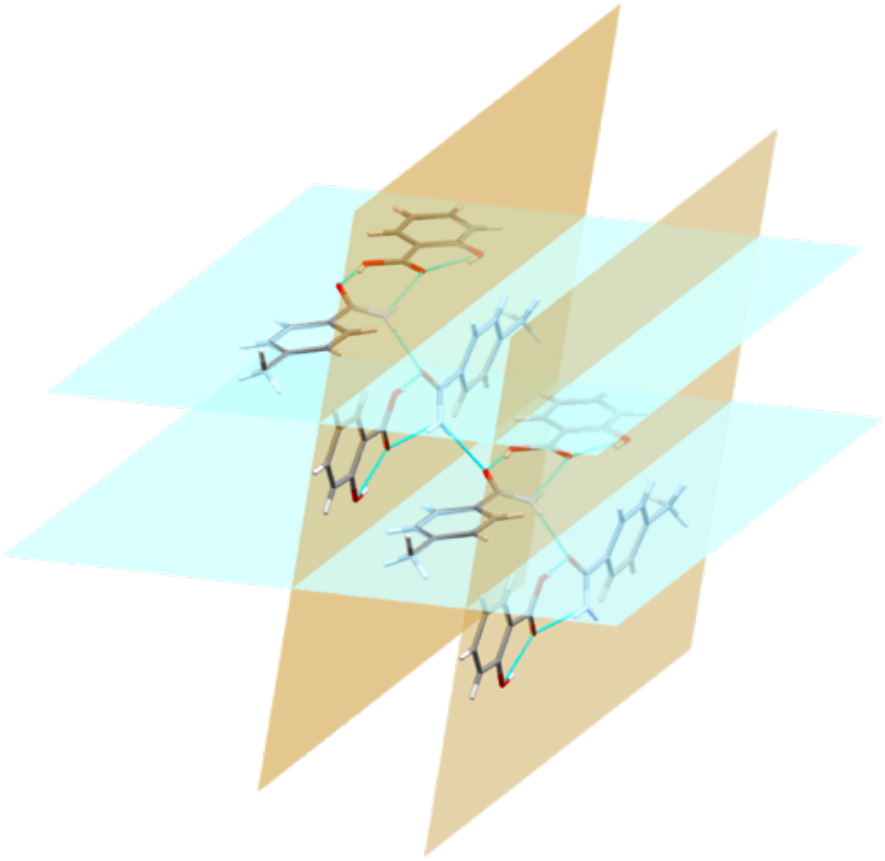


Figure 36: Alternating planar dimers along the *a*-axis in 60.66 cocrystal.

The second cocrystal of **60**, with 2-fluorobenzamide[**70**], crystallizes in the $P2_1/c$ space group [a = 23.894(3), b = 5.2201(7), c = 22.767(3), β = 114.340(3), V =

2587.3 Å^3]. The primary motif observed here is the robust $R_2^2(8)$ dimer as above, however, the overall architecture of this structure is decidedly different from the **60.66** cocrystal. Instead of the formation of a C(4) chain, the second amide proton participates in N-H···O-H hydrogen bonding interaction to the phenol group, forming a binary level $R_4^4(16)$ tetramer (Figure 37). These tetramers line up down the *b*-axis, however, there is no evidence of π···π stacking interactions observed, with the rings separated by approx. 5.2 Å and offset by 138° (Figure 38).

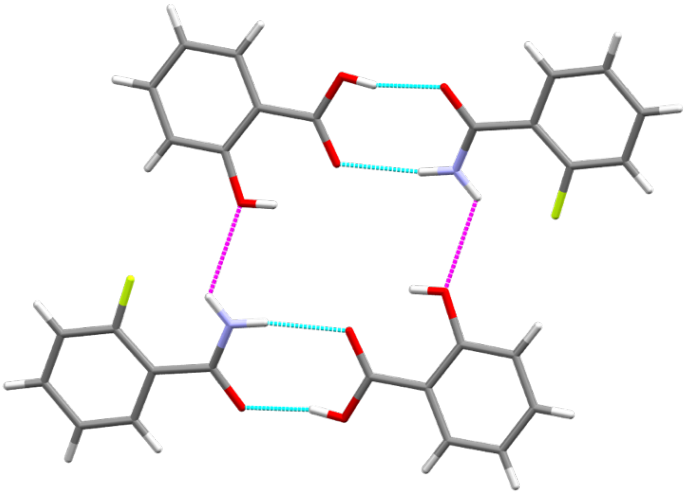


Figure 37: $R_4^4(16)$ tetramer formed in 60.70 cocrystal.

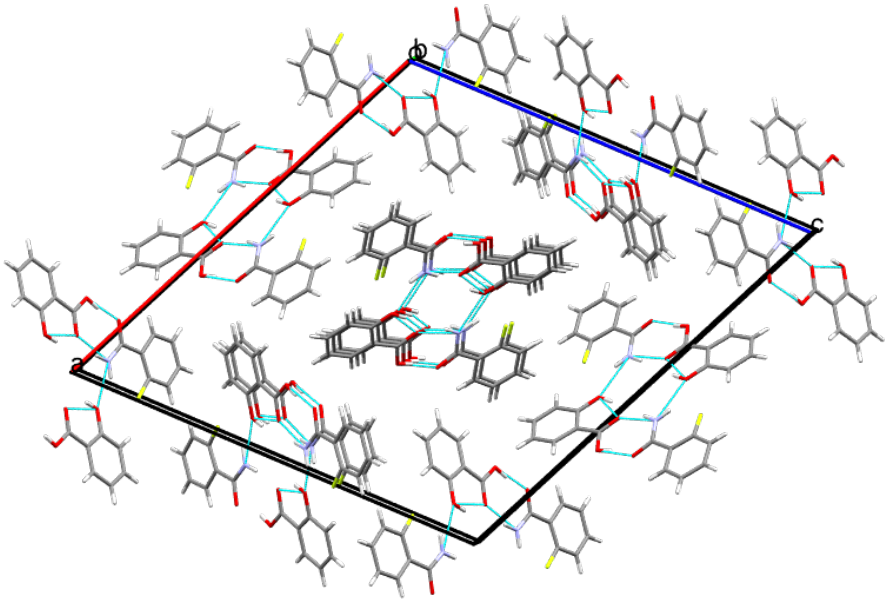


Figure 38: Overall packing observed in 60.70 cocrystal (view down *b*-axis).

3.6 Towards Understanding the Reactive Crystallizations of Salsalate

3.6.1 Reactive Cocrystallization

The hydrolysis and subsequent cocrystallization that was observed in screening for salsalate cocrystals above is a phenomenon that has not been previously described in the literature. In this context, we coined the term 'reactive cocrystallization' to describe this process, and our efforts turned to understanding the underlying mechanism behind the process.

In the cocrystallization screen of salsalate [**58**], all of the evidence for the formation of salicylic acid cocrystals and/or salicylic acid [**60**] occurred from the methanol solvent crystallizations. It was postulated that salsalate may potentially be unstable with respect to hydrolysis in methanol. In order to confirm this hypothesis, **58** was stirred in methanol for a period of 1 week, after which the methanol was removed *in vacuo* and the resultant solid investigated for the presence of **60** *via* NMR. There was no evidence of the formation of **60** in this case.

A second experiment was performed, by boiling the salsalate in methanol, from which salicylic acid was identified. This result, however, does not fully explain the results identified in the cocrystallization studies, since they were performed at room temperature, although it confirms that the salsalate is susceptible to hydrolysis in methanol under certain conditions.

Structural degradation of salsalate in contact with alkaline excipients was studied by Worn *et al*. in 1983.[36] The outcome of the study was that the presence of alkaline impurities would adversely affect the stability of salsalate. It is unlikely that the amides for which the reactive cocrystallization was observed are sufficiently basic to cause the hydrolysis to occur. Amides are very weakly basic, and unlikely to catalyse the hydrolysis of the salsalate. If the alkalinity of the amides were a factor, the instability of the salsalate would have been observed in more of the crystallization vials, and with more solvents, and so it is unlikely that the amides are chemically catalysing the hydrolysis of the salsalate.

The evidence indicated that some other factor must be at play in the cocrystallization experiments to cause the hydrolysis to occur. At this point, we decided to investigate the relative energies and interactions of the different salsalate forms from Hirshfeld surface analysis (HSA) and density functional theory (DFT).

HSA has demonstrated that there is a clear difference between the two polymorphs; in particular, the percentage contributions of the C-C, C-H, H-H and O-H bonds to the structures (Table 5). However, both display the closest molecular proximity at the carboxylic acid groups (Figure 39). Form II displays a lower percentage contribution of the O-H bonds, likely due to the lack of atropisomerism around C_7-C_8, and a far greater contribution of H-H bonds. The fingerprint plots for the molecules show a more localised interaction distance for Form II vs Form I (Figure 40). Hirshfeld surface analysis successfully demonstrated the differences between the two forms, however, it did not provide additional insight into the reasons for the observed reactive cocrystallization.

Table 5: Comparison of the different intermolecular contacts present in the two polymorphs of salsalate [58].

Bond	Form I Contribution (%)	Form II Contribution (%)
C-C	6.6%	9.2%
C-H	19.2%	14.3%
C-O	5.6%	7.8%
H-H	31.6%	39.6%
O-H	35.5%	27.6%
O-O	1.5%	1.6%

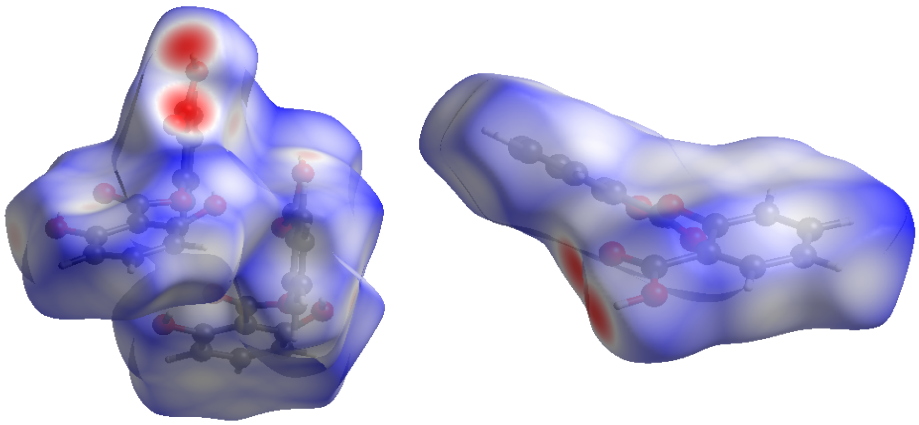


Figure 39: Hirshfeld surface diagrams for 58 Form I (left) and Form II (right), close proximity (red), moderate proximity (white), low proximity (blue) of other molecules in the crystal structure.

DFT calculations were performed to determine the relative lattice energies of the two forms. The calculations revealed that Form I has a lattice energy value of 360 kJ mol^{-1}, whereas Form II has a value of 225 kJ mol^{-1}. This confirms that Form II is the more stable of the two forms, as had been demonstrated in thermal analysis measurements.

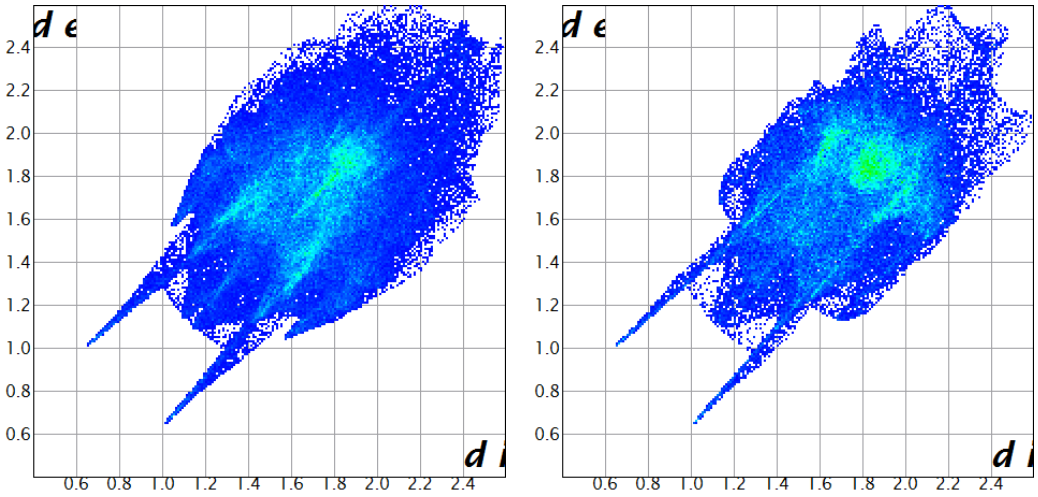


Figure 40: d_e vs d_i fingerprint plots for 58 Form I (left) and Form II (right).

At this point, the mechanistic pathway by which reactive cocrystallization occurs with **58** is unclear. One potential explanation for this is the reduced bond rotation around C_7-C_8 in Form II. Form II has only been observed in crystallizations from methanol, and reactive cocrystallization was observed in the same solvent. It is possible that the formation of the strong $R_2^2(8)$ dimer hydrogen bond with the carboxylic acid moiety, coupled with the reduced bond rotation, allows for hydrolytic cleavage of the salsalate molecules, a reaction that is prevented by the atropisomerism normally.

NMR experiments were undertaken to investigate the effects of the amides **66** and **70** upon the stability of **58** in methanol. The materials were dissolved in deuterated methanol in a 1:1 ratio and investigated *via* ^{1}H NMR once per day for 8 days. Evidence of structural changes were apparent by day 3, albeit at a low concentration, and the observed concentration increased over time (Figure 41), indicating that the reactive cocrystallization process is slow, which would agree with the stability of the material observed in the methanol experiments above.

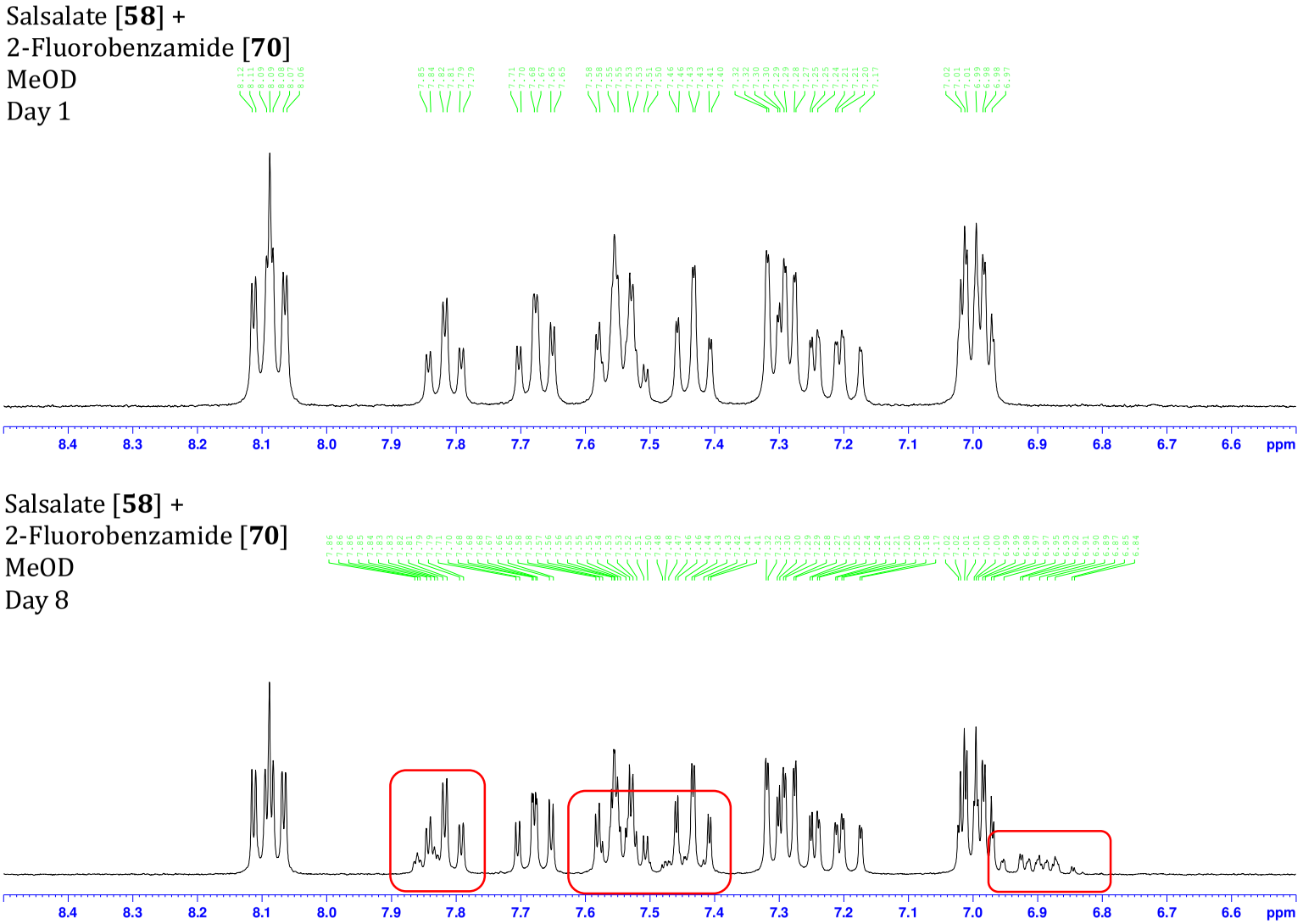


Figure 41: NMR comparison of 1:! mixture of 58.66 in MeOD, day 1 (top) vs day 8 (bottom), evidence of reaction circled in red.

3.6.2 Reactive Salt Formation

Examination of the solution crystallization behaviour of all coformers used in the second screen (Section 3.4) revealed reactive salt formation in two cases, one of which was novel. One of these involved 4-aminopyridine [**96**]. 4-Aminopyridine [**96**] is used to treat certain symptoms of multiple sclerosis,[37] and is sold to the European market under the tradename Fampyra. The salt structure determined in this study was consistent with that of 4-aminopyridinium salicylate [**106**], reported by Fun *et al*. in 2010 (Figure 42).[38]

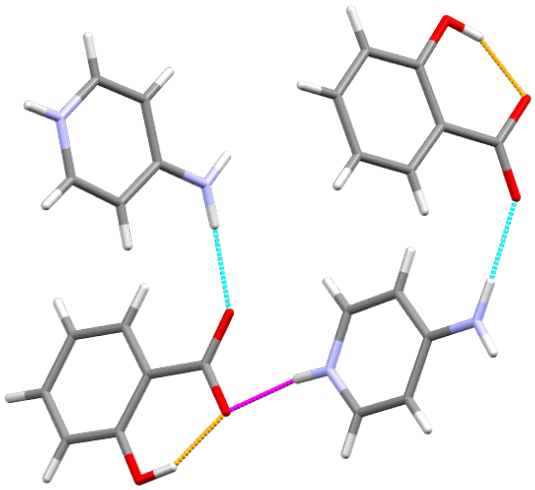


Figure 42: Structure of 4-aminopyridinium salicylate (1:1) [106] [DUSYOG].[38]

The other compound formed was 4-hydroxypyridinium salicylate [**107**]. This novel salt material crystallized in $P2_1/n$, and displays the similar intermolecular hydrogen bonding interactions to that observed in [**106**], above [a = 8.777(2) Å, b = 12.402(3) Å, c = 10.924(2) Å, β = 111.727(7)°, V = 1104.6(4) Å^3]. In both cases, the *para* group (hydroxyl or amino) participates in a hydrogen bond to the C=O not involved with the pyridyl group, and an S(6) ring is formed within the salicylate.

The salt formation in this case can be rationalised, given that both **96**, and 4-hydroxypyridine [**97**] can act as bases, catalysing the hydrolysis of the salsalate and forming the salicylic acid product, which subsequently protonates the pyridine forming the salt product. This is consistent with the instability of salsalate in the presence of alkaline impurities observed by Worn *et al*.[36]

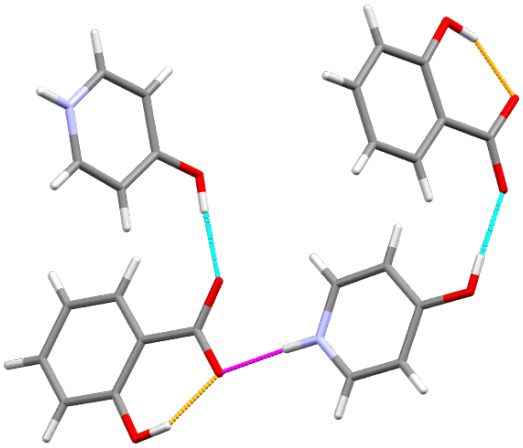


Figure 43: Intermolecular interactions observed in 1:1 salt of 4-hydroxypyridinium salicylate [107].

3.7 Conclusions

Investigations into the cocrystallization of salsalate led to the successful preparation of a previously unknown polymorphic form of salsalate, along with 8 novel cocrystals of salsalate [**58**], two novel cocrystals of salicylic acid, and one novel salt of salicylic acid. It is possible that the results generated in this study could have implications for salsalate formulations in the future, as the stability of this material in the presence of certain GRAS materials could be called into question. The exact rationale for the reactive cocrystallization observed here remains unclear, with further experimentation required to probe the mechanism and scope of this phenomenon.

Our low success rate in the initial 32-compound cocrystallization screen lead us to investigate alternative knowledge based methodologies to streamline the coformer selection process and gain a more accurate level of prediction prior to commencing cocrystallization screening in the future.

Chapter 3

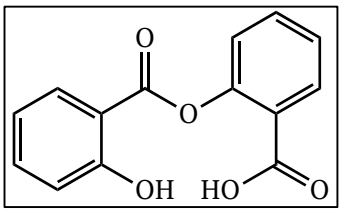
Experimental

3.8 Experimental

Preparation and characterisation of all materials in this experimental were performed in the same manner as outlined in general procedures (Section 2.7) of this thesis.

3.8.1 Salsalate Polymorph

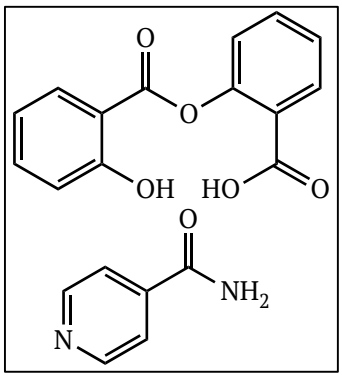
Salsalate [58] Form II



Colourless needles of **58** Form II were obtained from slow dissolution of **58** in methanol in a 12 mm sample vial, capped with just a pin-hole to allow solvent to escape. Crystals were recovered over 3-4 weeks; mp 147-149 °C; ν_{max} (ATR)/cm^{-1}; 3465 (w) (O-H stretch), 1676 (C=O), 1608, 1578 (Aromatic C=C bend), 1470 (Aromatic C=C bend), 1411, 1332, 1277, 1263, 1245 (C-O), 1155 (C-O), 1126 (C-O), 775, 747. Crystal data for **58** Form II: $C_{14}H_{10}O_5$, *Mr* = 258.22, monoclinic, $P2_1/n$, a = 16.397(4) Å, b = 4.0573(8) Å, c = 18.583(4) Å, β = 102.101(5) °, V = 1208.8(4) Å^3, Z = 4, D_c = 1.419 g cm^{-3}, F_{000} = 536, Mo Kα radiation, λ = 0.71073 Å, T = 296.(2) K, $2\theta_{max}$ = 25.12°, μ = 0.109 mm^{-1}, 13859 reflections collected, 2150 unique (R_{int} = 0.0698), final GooF = 0.833, R_1 = 0.0456 [1167 obs. data: $I>2\sigma(I)$], wR_2 = 0.1681 (all data).

3.8.2 Salsalate Cocrystals

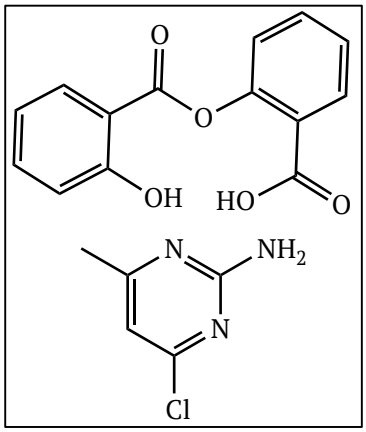
Salsalate isonicotinamide cocrystal [58.61]



2-((2-hydroxybenzoyl)oxy)benzoic acid [**58**] (0.0388 g, 0.25 mmol) and isonicotinamide [**61**] (0.0183 g, 0.25 mmol) were used. Colourless plate crystals of **58.61** were obtained from acetonitrile. DSC (endotherm): 126-128 °C; ν_{max} (ATR)/cm^{-1}; 3349 (w) (N-H stretch), 1681 (C=O), 1387, 1192 (C-O), 1152 (C-O), 1032, 748, 693, 667, 647. Crystal data for **58.61**:

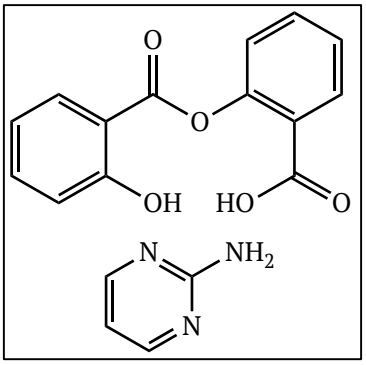
$C_{20}H_{16}N_2O_6$, *Mr* = 380.35, triclinic, *P-1*, *a* = 6.9660(12) Å, *b* = 7.1847(14) Å, *c* = 18.983(4) Å, α = 94.277(5) °, β = 95.816(4) °, γ = 104.860(4) °, *V* = 908.6(3) Å^3, *Z* = 2, D_c = 1.390 g cm^{-3}, F_{000} = 396, Mo Kα radiation, λ = 0.71073 Å, *T* = 296.(2) K, $2\theta_{max}$ = 25.81°, μ = 0.104 mm^{-1}, 26431 reflections collected, 3488 unique (R_{int} = 0.0395), final GooF = 1.021, R_1 = 0.0399 [2429 obs. data: $I>2\sigma(I)$], wR_2 = 0.1162 (all data).

Salsalate 2-amino-4-chloro-6-methylpyrimidine cocrystal [58.93]



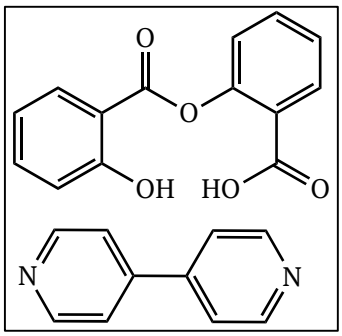
2-((2-hydroxybenzoyl)oxy)benzoic acid [**58**] (0.0776 g, 0.50 mmol) and 2-amino-4-chloro-6-methylpyrimidine [**93**] (0.0718 g, 0.50 mmol) were used. Colourless plate crystals of **58.93** were obtained from acetonitrile. DSC (endotherm): 137-141 °C; ν_{max} (ATR)/cm^{-1}; 3466, 3405, 3316 (w) (N-H stretches), 1738 (C=O), 1685 (C=O), 1576 (Aromatic C=C bend), 1554 (Aromatic C=C bend), 1293 (C-O), 1213 (C-O), 1196 (C-O), 1105 (C-O), 1087, 849, 757, 748 (C-Cl), 692. Crystal data for **58.93**: $C_{19}H_{16}ClN_3O_5$, *Mr* = 401.80, triclinic, *P-1*, *a* = 7.6518(3) Å, *b* = 9.4533(4) Å, *c* = 14.2495(6) Å, α = 83.969(2) °, β = 74.507(2) °, γ = 72.482(2) °, *V* = 947.07(7) Å^3, *Z* = 2, D_c = 1.409 g cm^{-3}, F_{000} = 416, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 67.03°, μ = 2.111 mm^{-1}, 12425 reflections collected, 3238 unique (R_{int} = 0.0208), final GooF = 0.921, R_1 = 0.0408 [3036 obs. data: $I>2\sigma(I)$], wR_2 = 0.1359 (all data).

Salsalate 2-aminopyrimidine 2:1 cocrystal [58.95]



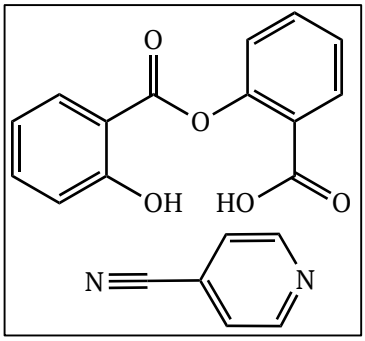
2-((2-hydroxybenzoyl)oxy)benzoic acid [**58**] (0.0776 g, 0.50 mmol) and 2-aminopyrimidine [**95**] (0.0238 g, 0.25 mmol) were used. DSC (endotherm): 105-107 °C; ν_{max} (ATR)/cm^{-1}; 3332, 3163 (w) (N-H stretches), 1670 (C=O), 1650 (C=O), 1577 (Aromatic C=C bend), 1557 (Aromatic C=C bend), 1476 (Aromatic C=C bend), 1338, 1208 (C-O), 1180 (C-O), 1156 (C-O), 1128 (C-O), 1086, 1066, 787, 710.

Salsalate 4,4'-bipyridyl cocrystal [58.15]



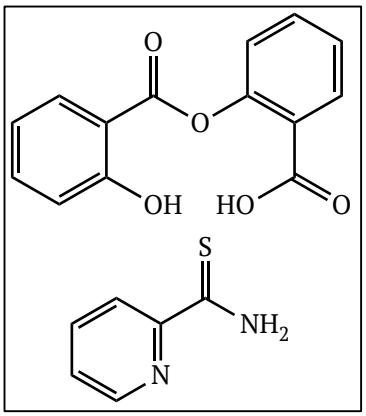
2-((2-hydroxybenzoyl)oxy)benzoic acid [**58**] (0.0776 g, 0.50 mmol) and 4,4'-bipyridyl [**15**] (0.0781 g, 0.50 mmol) were used. Colourless plate crystals of **58**.**15** were obtained from acetonitrile. DSC (endotherm): 61-66 °C; ν_{max} (ATR)/cm^{-1}; 1687 (C=O), 1263, 1244 (C-O), 1181 (C-O), 1154 (C-O), 1126 (C-O), 1084, 806, 751. Crystal data for **58**.**15**: $C_{19}H_{14}NO_5$, *Mr* = 380.35, triclinic, *P-1*, *a* = 7.6250(12) Å, *b* = 7.9985(11) Å, *c* = 13.368(2) Å, α = 95.770(5) °, β = 97.757(5)°, γ = 94.778(5) °, *V* = 799.9(2) Å^3, *Z* = 2, D_c = 1.396 g cm^{-3}, F_{000} = 350, Mo Kα radiation, λ = 0.71073 Å, *T* = 300(2) K, $2\theta_{max}$ = 26.37°, μ = 0.102 mm^{-1}, 8782 reflections collected, 3219 unique (R_{int} = 0.0369), final GooF = 1.025, R_1 = 0.0417 [2158 obs. data: $I>2\sigma(I)$], wR_2 = 0.1425 (all data).

Salsalate 4-pyridinecarbonitrile cocrystal [58.100]



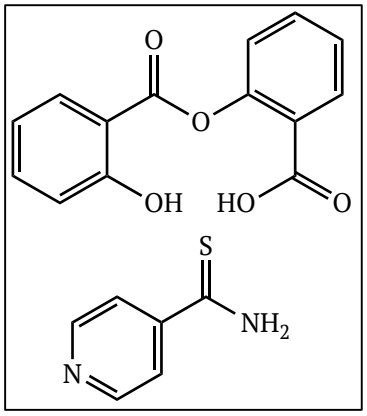
2-((2-hydroxybenzoyl)oxy)benzoic acid [**58**] (0.0776 g, 0.50 mmol) and 4-pyridinecarbonitrile [**100**] (0.0521 g, 0.50 mmol) were used. Colourless plate crystals of **58.100** were obtained from isopropanol. DSC (endotherm): 102-104 °C; ν_{max} (ATR)/cm^{-1}; 2242 (w) (C≡N) 1681 (C=O), 1291, 1247 (C-O), 1186 (C-O), 799, 750. Crystal data for **58.100**: $C_{20}H_{14}N_2O_5$, *Mr* = 362.33, monoclinic, *Pc*, *a* = 15.7904(14) Å, *b* = 7.2513(6) Å, *c* = 31.105(3) Å, β = 100.760(2) °, *V* = 3498.9(5) Å^3, *Z* = 8, D_c = 1.376 g cm^{-3}, F_{000} = 1504, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 25.94°, μ = 0.101 mm^{-1}, 50412 reflections collected, 13486 unique (R_{int} = 0.0994), final GooF = 1.060, R_1 = 0.1085 [6999 obs. data: $I>2\sigma(I)$], wR_2 = 0.3660 (all data).

Salsalate 2-pyridinethioamide cocrystal [58.102]



2-((2-hydroxybenzoyl)oxy)benzoic acid [**58**] (0.0776 g, 0.50 mmol) and 2-pyridinethioamide [**102**] (0.0691 g, 0.50 mmol) were used.. DSC (endotherm): 110-113 °C; ν_{max} (ATR)/cm^{-1}; 3468, 3346 (w) (N-H stretches), 1680 (C=O), 1603, 1579 (Aromatic C=C bend), 1333, 1301, 1276, 1263, 1245 (C-O), 1196 (C-O), 1128 (C-O), 1105 (C-O), 797, 749.

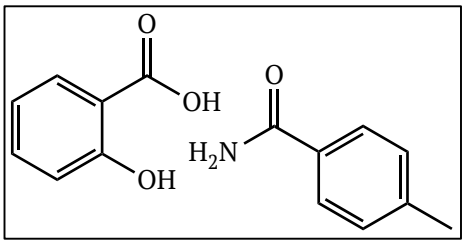
Salsalate 4-pyridinethioamide cocrystal [58.103]



2-((2-hydroxybenzoyl)oxy)benzoic acid [**58**] (0.0776 g, 0.50 mmol) and 4-pyridinethioamide [**103**] (0.0691 g, 0.50 mmol) were used. DSC (endotherm): 123-127 °C; ν_{max} (ATR)/cm^{-1}; 3468, 3325 (w) (N-H stretches), 1678 (C=O), 1413, 1301, 1279, 1264, 1246 (C-O), 1155 (C-O), 1128 (C-O), 1065, 925 (C=S), 749.

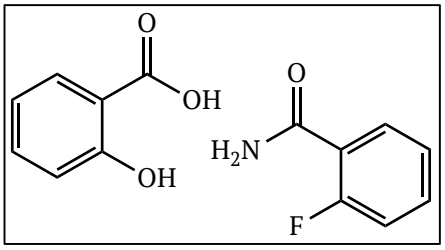
3.8.3 Salicylic Acid Cocrystals

Salicylic acid 4-methylbenzamide cocrystal [60.66]



2-((2-hydroxybenzoyl)oxy)benzoic acid [**58**] (0.0388 g, 0.25 mmol) and 4-methylbenzamide [**66**] (0.0338 g, 0.25 mmol) were used. Colourless plate crystals of **60.66** were obtained from methanol; mp 107-109 °C; ν_{max} (ATR)/cm^{-1}; 3469 (w) (O-H stretch), 3340 (w) (N-H stretch), 1673 (C=O), 1614 (C=O), 1449 (Aromatic C=C bend), 1412, 1156, 1125, 692, 652, 631. Crystal data for **60.66**: $C_{15}H_{15}NO_4$, *Mr* = 273.28, orthorhombic, *Pbca*, *a* = 9.9648(9) Å, *b* = 8.3061(8) Å, *c* = 33.582(3) Å, *V* = 2779.5(5) Å^3, *Z* = 8, D_c = 1.306 g cm^{-3}, F_{000} = 1152, Mo Kα radiation, λ = 0.71073 Å, *T* = 300(2) K, $2\theta_{max}$ = 22.29°, μ = 0.095 mm^{-1}, 22445 reflections collected, 1765 unique (R_{int} = 0.0457), final GooF = 1.029, R_1 = 0.0370 [1364 obs. data: $I>2\sigma(I)$], wR_2 = 0.1093 (all data).

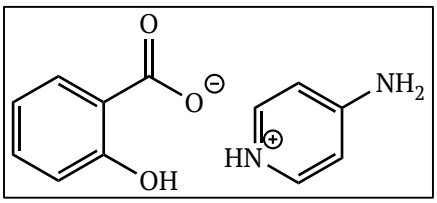
Salicylic acid 2-fluorobenzamide cocrystal [60.70]



2-((2-hydroxybenzoyl)oxy)benzoic acid [**58**] (0.0388 g, 0.25 mmol) and 2-fluorobenzamide [**70**] (0.0348 g, 0.25 mmol) were used. Colourless block crystals of **60.70** were obtained from methanol; mp : 92-94 °C; ν_{max} (ATR)/cm^{-1}; 3469 (w) (O-H stretch), 3390 (w) (N-H stretch), 1645 (C=O), 1614 (C=O), 1404, 1300, 757, 749. Crystal data for **60.70**: $C_{14}H_{12}FNO_4$, *Mr* = 277.25, monoclinic, *P2₁/c*, *a* = 23.894(3) Å, *b* = 5.2201(7) Å, *c* = 22.767(3) Å, *β* = 114.340(3)°, *V* = 2587.3(6) Å^3, *Z* = 8, D_c = 1.424 g cm^{-3}, F_{000} = 1152, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 26.70°, μ = 0.114 mm^{-1}, 37778 reflections collected, 5436 unique (R_{int} = 0.0693), final GooF = 1.061, R_1 = 0.0886 [3399 obs. data: $I>2\sigma(I)$], wR_2 = 0.2569 (all data).

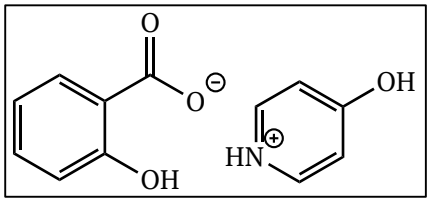
3.8.4 Salicylic Acid Salts

4-Aminopyridinium salicylate [106]



2-((2-hydroxybenzoyl)oxy)benzoic acid [**58**] (0.0195 g, 0.25 mmol) and 4-aminopyridine [**96**] (0.0471 g, 0.50 mmol) were used. Colourless plate crystals of **106** were obtained from methanol; DSC (endotherm): 182-184 °C; ν_{max} (ATR)/cm^{-1}; 3416, 3314 (w) (N-H stretches), 1632 (C=O), 1572 (Aromatic C=C bend), 1524 (Aromatic C=C bend), 1485 (Aromatic C=C bend), 1459 (Aromatic C=C bend), 1384, 1333, 1188 (C-O), 1140 (C-O), 857, 834, 810, 769. Crystal data for **106**: $C_{12}H_{12}N_2O_3$, *Mr* = 232.24, orthorhombic, *Pbca*, *a* = 12.242(3) Å, *b* = 11.527(3) Å, *c* = 16.541(5) Å, *V* = 2334.2(11) Å^3, *Z* = 8, D_c = 1.322 g cm^{-3}, F_{000} = 976, Mo Kα radiation, λ = 0.71073 Å, *T* = 300(2) K, $2\theta_{max}$ =26.46°, μ = 0.097 mm^{-1}, 47321 reflections collected, 2400 unique (R_{int} = 0.0832), final GooF = 1.074, R_1 = 0.0493 [1497 obs. data: $I>2\sigma(I)$], wR_2 = 0.0863 (all data).

4-Hydroxyopyridinium salicylate [107]



2-((2-hydroxybenzoyl)oxy)benzoic acid [**58**] (0.0195 g, 0.25 mmol) and 4-hydroxypyridine [**107**] (0.0476 g, 0.50 mmol) were used. Colourless plate crystals of **107** were obtained from methanol; mp 87-90 °C; ν_{max} (ATR)/cm^{-1}; 3467 (w) (N-H stretch), 1680 (C=O), 1481 (Aromatic C=C bend), 1190 (C-O), 1155 (C-O), 1085, 1065, 748. Crystal data for **107**: $C_{12}H_{11}NO_4$, Mr = 233.22, monoclinic, $P2_1/n$, a = 8.777(2) Å, b = 12.402(3) Å, c = 10.924(2) Å, V = 1104.6(4) Å^3, Z = 4, D_c = 1.402 g cm^{-3}, F_{000} = 488, Mo Kα radiation, λ = 0.71073 Å, T = 300(2) K, $2\theta_{max}$ =26.49°, μ = 0.107 mm^{-1}, 6591 reflections collected, 2242 unique (R_{int} = 0.0379), final GooF = 0.850, R_1 = 0.0442 [1412 obs. data: I>2σ(I)], wR_2 = 0.01593 (all data).

Chapter 3

References

3.9 References

(1) Anderson, K.; Wherle, L.; Park, M.; Nelson, K.; Nguyen, L. D. *Am. Heal. Drug Benefits* **2014**, *7*, 231–235.

(2) Jack, D. B. *Lancet* **1997**, *350*, 437–439.

(3) Cox, P. J.; Gilmour, G. I.; MacManus, S. M. *Int. J. Pharm.* **2000**, *204*, 133–136.

(4) Rainsford, K. D.; Buchanan, W. W. *Baillieres. Clin. Rheumatol.* **1990**, *4*, 247–268.

(5) Lanas, A. *Curr. Treat. Options Gastroenterol.* **2006**, *9*, 147–156.

(6) Scheiman, J. M.; Elta, G. H. *Semin. Arthritis Rheum.* **1990**, *20*, 121–127.

(7) Rainsford, K. D. *Aspirin and Related Drugs*, 1st ed.; Taylor and Francis, New York, 2004.

(8) Goldfine, A. B.; Fonseca, V.; Jablonski, K. A.; Chen, Y. I.; Tipton, L.; Staten, M. A.; Shoelson, S. E. *Ann. Intern. Med.* **2014**, *159*, 1–12.

(9) Min, S.-W.; Chen, X.; Tracy, T. E.; Li, Y.; Zhou, Y.; Wang, C.; Shirakawa, K.; Minami, S. S.; Defensor, E.; Mok, S. A.; Sohn, P. D.; Schilling, B.; Cong, X.; Ellerby, L.; Gibson, B. W.; Johnson, J.; Krogan, N.; Shamloo, M.; Gestwicki, J.; Masliah, E.; Verdin, E.; Gan, L. *Nat. Med.* **2015**, *21*, 1154–1162.

(10) R., G. C.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. *Acta Crystallogr. Sect. B* **2016**, *72*, 171–179.

(11) Greener, B.; Archibald, S. J.; Hodkinson, M. *Agnew. Chem. Int. Ed.* **2000**, *39*, 3601–3604.

(12) Etter, M. C. *J. Phys. Chem.* **1991**, *95*, 4601–4610.

(13) Ach, L.; Sutter, T. Salicylic acid derivative US922995, May 25, 1909.

(14) Anschutz, L.; Neher, R.; Kohlrausch, K. W. F. *J. fuer Prakt. Chemie* **1941**, *159*, 264–272.

(15) Einhorn, A. *Berichte der Dtsch. Chem. Gesellschaft* **1911**, *44*, 431–440.

(16) Habgood, M. *Cryst. Growth Des.* **2013**, *13*, 4549–4558.

(17) Wood, P. A.; Feeder, N.; Furlow, M.; Galek, P. T. A.; Groom, C. R.; Pidcock, E. *CrystEngComm* **2014**, *16*, 5839–5848.

(18) Vishweshwar, P.; McMahon, J. A.; Bis, J. A.; Zaworotko, M. J. *J. Pharm. Sci.* **2006**, *95*, 499–516.

(19) Fukte, S. R.; Wagh, M. P.; Rawat, S. *Int. J. Pharm. Pharm. Sci.* **2014**, *6*, 9–14.

(20) Hunter, C. A. *Angew. Chemie - Int. Ed.* **2004**, *43*, 5310–5324.

(21) Shattock, T. R.; Arora, K. K.; Vishweshwar, P.; Zaworotko, M. J. *Cryst. Growth Des.* **2008**, *8*, 4533–4545.

(22) McMahon, J. A.; Bis, J. A.; Vishweshwar, P.; Shattock, T. R.; McLaughlin, O. L.; Zaworotko, M. J. *Z. fur Krist.* **2005**, *220*, 340–350.

(23) Zhou, Z.; Chan, H. M.; Sung, H. H. Y.; Tong, H. H. Y.; Zheng, Y. *Pharm. Res.* **2016**, *33*, 1030–1039.

(24) Aakeroy, C. B.; Beatty, A. M.; Helfrich, B. A. *J. Am. Chem. Soc.* **2002**, *124*, 14425–14432.

(25) Eccles, K. S.; Elcoate, C. J.; Maguire, A. R.; Lawrence, S. E. *Cryst. Growth Des.* **2011**, *11*, 4433–4439.

(26) Perlovich, G. L. *CrystEngComm* **2015**, *17*, 7019–7028.

(27) Shulga, O.; Dunn, J. *Thermochim. Acta* **2004**, *410*, 15–21.

(28) Ebenezer, S.; Muthiah, P. T.; Butcher, R. J. *Cryst. Growth Des.* **2011**, *11*, 3579–3592.

(29) Macrae, C. F.; Bruno, I. J.; Chisholm, J. A.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; van de Streek, J.; Wood, P. A. J. *J. Appl. Crystallogr.* **2008**, *41*, 466–470.

(30) Johnson, T. B.; Sprague, J. M. *J. Am. Chem. Soc.* **1938**, *60*, 1622–1624.

(31) Jensen, K. A.; Falkenberg, P.; Thorsteinsson, T.; Lauridsen, M. *Dan. Tidsskr. Farm.* **1942**, *16*, 141–153.

(32) Wibaut, J. P.; Dingemanse, E.; Wilbaut, J. P.; Dingemanse, E. *Recl. des Trav. Chim. des Pays-Bas la Belgique* **1923**, *42*, 240–250.

(33) Poziomek, E. J.; Kramer, D. N.; Mosher, W. A. *J. Org. Chem.* **1960**, *25*, 2135–2137.

(34) Karrer, P.; Schukri, J. *Helv. Chim. Acta* **1945**, *28*, 820–824.

(35) McMillan, F. H.; Leonard, F.; Meltzer, R. I.; King, J. A. *J. Am. Pharm. Assoc.* **1953**, *42*, 457–464.

(36) Worn, A. L.; Mruw, P. O. P.; Irwin, W. J. *Int. J. Pharm.* **1983**, *16*, 115–123.

(37) Jensen, H. B.; Ravnborg, M.; Dalgas, U.; Stenager, E. *Ther. Adv. Neurol. Disord.* **2014**, *7*, 97–113.

(38) Fun, H. K.; Hemamalini, M.; Rajakannan, V. *Acta Crystallogr. Sect. E Struct. Reports Online* **2010**, *66*, o2010–o2011.

Chapter 4

Cocrystallization Prediction using Computational Methods

4.1 Introduction to Cocrystal Prediction

4.1.1 'Will it Crystallise?'

In late 2014, Jerome Wicker and Richard Cooper of the University of Oxford, published a report entitled 'Will it crystallize? predicting crystallinity of molecular materials'.[1] This paper outlined efforts to develop and test a machine learning tool to predict the crystallization propensity of organic materials using only their two-dimensional structures.

The conclusions of this work were very promising, the machine learning software was successful in training a classification algorithm that could predict whether or not an organic material would form crystals large enough for single crystal analysis with an impressive accuracy of 90.3%. The paper suggested that a potential application of the work would be to assess how functional group changes in a molecule could influence the crystallinity of a material, providing guidance for the crystal engineering process in drug design.

4.1.2 A Brief Overview of Machine Learning

'Machine learning' is the term used to describe the field of artificial intelligence that governs efforts to develop computational capacity for learning, without explicit programming of such abilities. The term 'machine learning' was coined in 1959 by Arthur Samuel, who developed the 'Samuels checkers-playing programme' – known to be the first example of a self-learning computer programme.[2]

Machine learning can be broadly classified into two subsections, each based on the type of learning concerned. Unsupervised learning attempts to find patterns or links in unlabelled sets of input data, whereas supervised learning, which is more commonly used, is based upon training the algorithm with a set of already classified predictor variables.[3] The supervised learning algorithm is trained by utilising a subset of the input 'training' dataset to make predictions, thereby mapping the best function to define the relationship between two related variables X and Y. X can be referred to as the input or 'predictor' variable, and Y

the output or 'response' variable.[4] Supervised learning methodologies can be further classified into 'regression' or 'classification' methods. Regression produces outputs that are continuous, such as a weight, where classification defines the output variables into one of two or more defined classes.[5]

The work of Wicker and Cooper used the support vector machine (SVM) as their algorithm of choice. The concept of SVM was introduced by Vapnik and Cortes for the case of 'optimal hyperplanes for separable classes'.[6] SVMs have been shown to provide a high level of performance when used for classification problems.[3] In a binary classification problem, the SVM is provided with data that falls into two categories (1 or 0), for which a hyperplane must be constructed that separates the data into either category [Figure 1]. The 'optimal' hyperplane is considered to be that which maximizes the margin (z) between the two closest data points to the hyperplane, thereby keeping the elements of the different classes furthest apart.

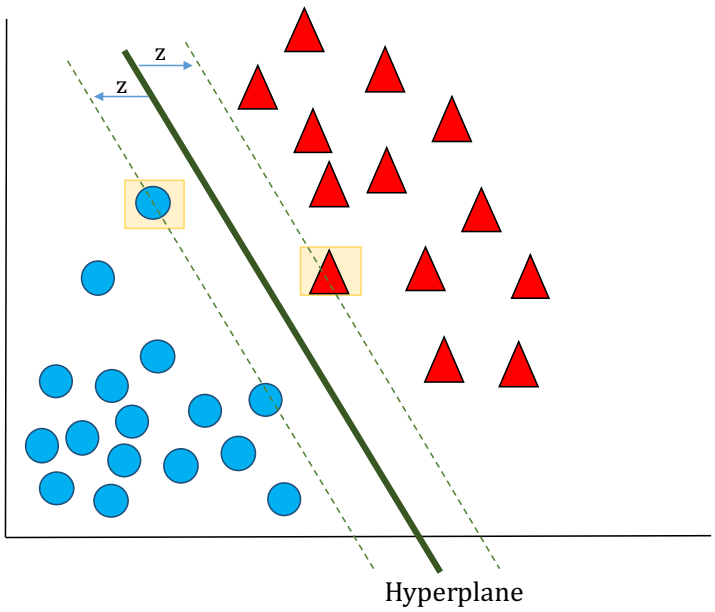


Figure 1: Example of a binary classification problem. The optimal hyperplane (green) is constructed such that the widest margin is achieved between the two classes (circles and triangles) [adapted from reference].[6]

In order to generate the optimal hyperplane necessary to classify whether or not a particular material would form crystals large enough for SCXRD, Wicker and Cooper used the SciKit learn package,[7] a python module integrating machine learning algorithms for use in both supervised and unsupervised methods. The scikit-learn package provides advantages over its competitors in terms of

computational efficiency and ease of use, with an accompanying user guide available online. In order to generate molecular descriptors for each molecule used in the training set, the RDKit cheminformatics toolkit was used.[8] This generates a large library of molecular descriptors for a given 2-dimensional molecular structure, incorporating both chemical and non-chemical functional information, which can be used to determine the parameters that best describe the optimal hyperplane for the given classification problem.

4.1.3 Computational Methods in Cocrystal Prediction

In the design and prediction of multi-component forms, the most commonly used methodology to anticipate/target the intermolecular interactions involved is still searching of the CSD,[9] followed by interpretation of the more common supramolecular synthons as a basis for coformer selection.[10] A computational prediction method based on the molecules themselves would provide significant benefits over a literature based method such as the above, allowing us to consider a broader range of functional groups for cocrystal screening, while also providing guidance as to the functional groups most likely to provide successful outcomes.

As discussed in section 1.1.1, efforts in crystal structure prediction research have advanced the field greatly, with very promising results in the most recent CCDC blind test.[11] In their 2008 report, Issa *et al*. leveraged the ongoing advancement in the field as a basis for the prediction of the structure 'of the most stable cocrystal of two given molecules in a fixed stoichiometry'.[12] The method for prediction proposed in the 2008 report was *ab initio* comparison of the lattice energies of the individual components of a cocrystal to that of the cocrystal itself. The study remarked that the differences between the relative energies of many of the cocrystals studied were insufficiently large as to allow for accurate cocrystallization prediction, but that the outcome of the method justified further work.

In 2009, Fabian[13] published an interesting statistical analysis of the CSD[9] in which he determined correlations between different qualitative aspects of molecules and their formation of cocrystals. The conclusions of the work

identified correlations, but each with significant exceptions and therefore, no steadfast rules for the formation of cocrystals. The strongest correlation identified was molecular polarity, with descriptors such as dipole moment and fractional polar volume appearing to provide particularly strong influence on cocrystal formation. In other words, molecules are most likely to form cocrystals with coformers of similar polarities to themselves. Secondly, molecular shape descriptors relating to the van der Waals volume identified a correlation that molecules tend to cocrystallize with coformers of similar shape and size to themselves. The third and final correlation identified in this work linked the formation of cocrystals to the strength of the potential hydrogen bonds that could form between them, which must be considered separately to the absolute number of hydrogen bond donors and acceptors.

In 2014, Grecu *et al*. published a cocrystal prediction methodology using molecular electrostatic potentials (MEPs) to compare the energy of a cocrystal to that of the pure coformers.[14] The approach used similar calculated values to those in Hunter's table,[15] to rank the presence of positive and negative surface site interaction points (SSIPs) on the molecules. The differences between the interaction energies for a ranked list of these SSIPs to that of the pure components provided an insight into the 'thermodynamic driving force for cocrystal formation'. Overall, the method performed very well when tested against the results of experimental cocrystallization screens; in general, materials which formed cocrystals were correctly assigned a large energy difference and ranked high in the list. The method compared well with the COSMO-RS method that uses excess enthalpy to estimate cocrystal formation, where it is assumed that the interactions occurring in the cocrystal are similar to that occurring in a supercooled liquid.[16]

Hydrogen bond propensity (HBP) calculations have been used successfully to predict the formation of cocrystalline systems.[10,17] HBP calculations, developed by Galek *et al*.,[18] describe donor and acceptor hydrogen bond relationships between molecular pairs. The HBP model gave correct classification of approximately 90% for 'sample donor and acceptor pairings' from a set of 1083 crystal structures. Wood *et al*.[10] have analysed the use of this method to predict

cocrystals, indicating a good level of success in correctly classifying the formation of these materials. Similarly, Delori *et al*.[17] used HBP calculations to aid the synthesis of cocrystals of pyrimethamine, suggesting that the method 'may be a useful tool in designing targeted screening experiments'. A HBP calculation tool is now built into the most recent version of Mercury structure visualisation software.[19]

4.2 Project Objectives

Having achieved an unexpectedly low number of successful outcomes in cocrystallization screening of salsalate utilising the traditional knowledge-based coformer selection techniques[10] (Chapter 3), it was envisaged that the machine learning approach could be used to predictively classify the formation (or not) of a cocrystal. A successful classification method would predict whether or not a cocrystal would form given only the molecular structures of the coformers, thereby reducing the number of unsuccessful outcomes in a cocrystallization screen, and concurrently, reducing the time and resource requirements for development of libraries of these novel materials.

The overall objectives of the project were:

(a) To develop a matrix of successful and unsuccessful cocrystallization experiments.

(b) To use this library in the training of a suitable machine learning algorithm with the aim of developing an effective cocrystallization prediction tool.

This research was conducted in collaboration with Prof. Richard Cooper and Jerome Wicker of the University of Oxford, Oxford, UK.

4.3 Experimental Data Matrix

In order to generate the binary data matrix required for SVM classification, knowledge of successful and unsuccessful cocrystallization screening experiments was required. Using traditional knowledge-based coformer selection techniques,[10] a matrix of 720 cocrystal combinations was designed and

split into four quadrants of 180 combinations each (Figure 2). The matrix was split into the quadrants to generate a manageable workload that could be split between UCC and Oxford. This matrix consisted of 20 coformers with 18 substituted benzoic acids and 18 substituted benzamides, some of which were already known in the literature[21-64] (Figure 3 and 4).

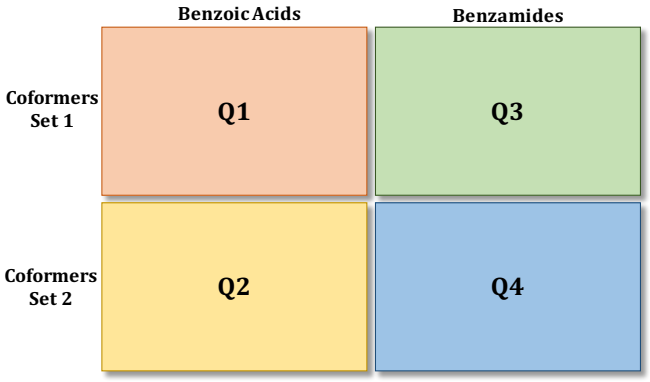


Figure 2: Summarizing how the experimental work was divided into four quadrants.

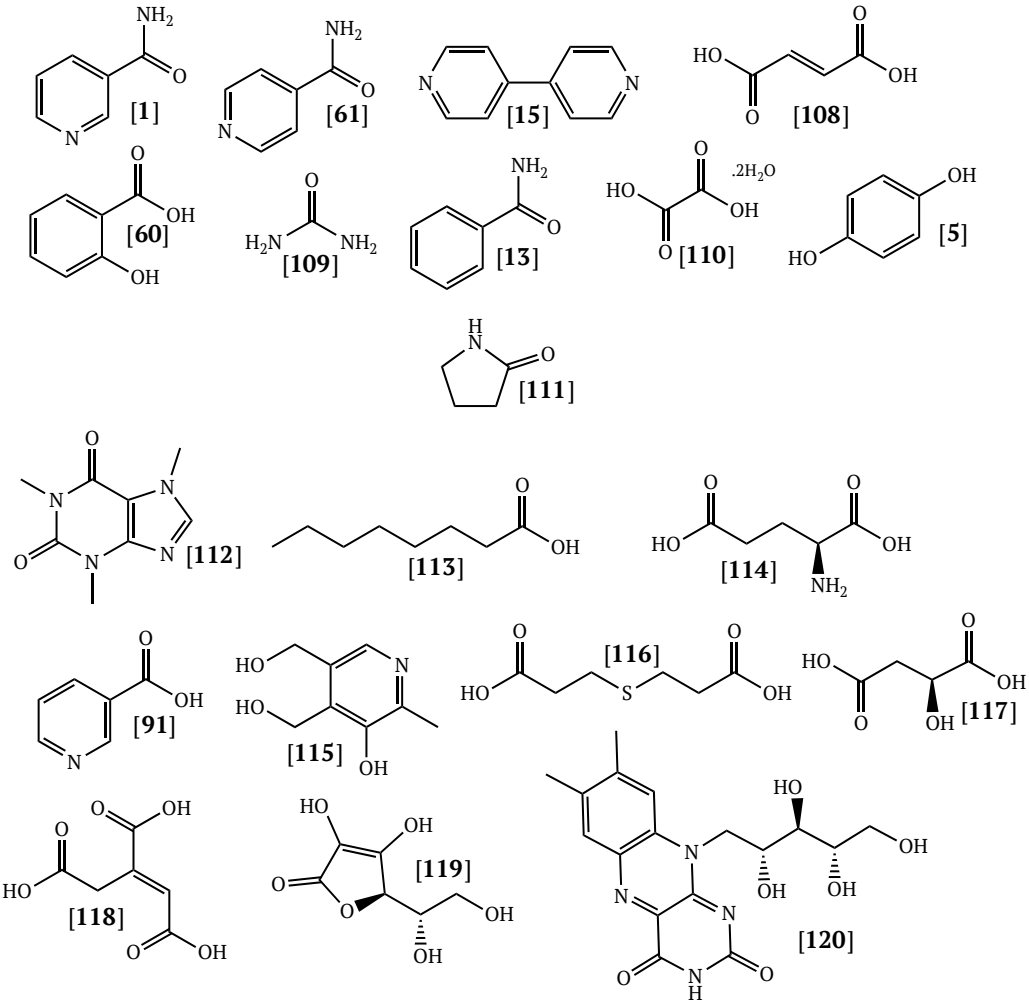


Figure 3: Coformer selection used for data matrix; set 1 (Q1 and Q2, top) and set 2 (Q3 and Q4, bottom).

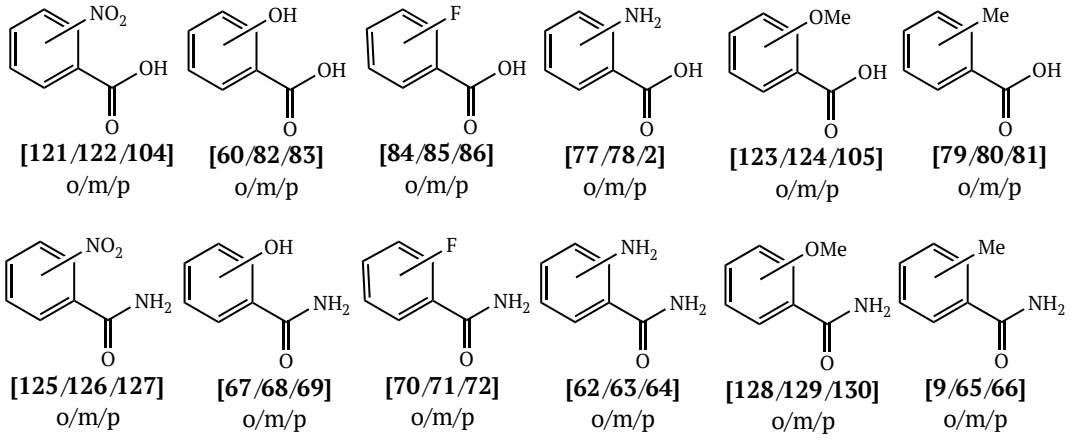


Figure 4: Substituted aromatic acids and amides used for intial matrix preparation (o/m/p refers to *ortho meta* or *para* substitution on the aromatic ring).

For the combinations not available in the literature, experimental cocrystallization screening was required to generate both the positive and negative results required for algorithm training. Q1 and Q2 were completed in this study in University College Cork, with subsequent completion of Q3 by Oliver Robshaw, and Q4 by Edmund Little, as part of their individual part II thesis studies at the University of Oxford.

Cocrystallization screening was undertaken using ball mill grinding of each coformer combination in a 1:1 stoichiometric ratio (as per Section 2.7.1), followed by analysis *via* PXRD and IR. Neat grinding was chosen for cocrystal screening in this case due to its efficiency in high-throughput screening and its high success rate in the successful formation of cocrystals.[20] Analysis *via* SCXRD was not chosen for this screening due to the large amount of time required to complete the structure solutions to confirm each novel material formed, and the difficulties surrounding preparation of suitable single crystals. Compounds **61**, **82** and **84** were not available for screening during the completion of Q3 and Q4 and so were omitted from Q3 and Q4. A small number of experiments which produced pastes were also omitted (assigned a value of '2'), leaving the final number of data points in the matrix at 657. The final data matrix contained 403 unsuccessful and 254 successful results. The successful results were spread across all four quadrants (85 in Q1, 42 in Q2, 84 in Q3, and 43 in Q4).

Salt formation was considered as a potential source of inaccuracy in the data matrix. However, since the coformers selected for screening contain relatively

few sites available for protonation, the projected number of potential salts was small. Therefore, for the purposes of the data matrix, there was no discrimination between cocrystals and salt, with any new solid form designated as '1'. The binary data matrix, including the cocrystals reported in the literature, are shown in Tables 1-4. Examples of PXRD and IR patterns for a successful cocrystallization experiment using nicotinamide [**1**] and 4-fluorobenzoic acid [**86**] are shown in Figures 5-7.

Table 1: Experimental results and literature results for cocrystal combinations in Quadrant 1 (1 indicates formation of a new solid form, 0 indicates no evidence of change observed in PXRD or IR).

Compound Name / Acid R-Group	Nicotinamide	Isonicotinamide	4,4'-Bipyridyl	Fumaric Acid	Salicylic Acid	Urea	Benzamide	Oxalic Acid Dihydrate	Hydro-quinone	2-Pyrrolid-inone
2-Nitro	1	1	GAWJEU[21]	0	0	1	1	0	0	1
3-Nitro	0	ASAXOH[22]	PUJHUY[23]	0	0	1	OVEZUL[24]	0	0	1
4-Nitro	1	AJAKEB[25]	DAQZIF/01[26,27]	0	0	1	YOPCAI[28]	0	0	LATREF[29]
2-Hydroxy	SODDOF/01[30,31] SODDOF02[32]	XAQQEM[33] QAFTID[34]	KONZEU[35] KOPKEH[35]	0	-	SLCDAC/01[36,37] SLCADC10[38]	URISAQ/01[24,39] QAFQEW[34]	0	0	1[b]
3-Hydroxy	XAQQIQ[33]	LUNMEM[40]	HONVAI/01[41,42]	1	0	1	0	1	1	1[b]
4-Hydroxy	RUYHEZ/01[43,44]	VAKTOR[45]	EPUPUB/01[46] EPUQEM[46]	0	0	1	0	0	0	1[b]
2-Fluoro	1	HANHEL[47]	1	0	0	1	1	0	0	0[b]
3-Fluoro	0	CACGUK[47]	0	0	0		1	0	0	0[b]
4-Fluoro	1	ASAXUN/01[22,47]	0	0	0	0	1	0	0	1[b]
2-Amino	1	1	NINXOZ[48]	0	RUTGUK[49]	0	0	1	0	1[b]
3-Amino	0	1	LEGPIY[50] UDUZIC[51]	1	1	0	0	MUGBIA[a,52]	0	0[b]
4-Amino	ABULIU[53]	SOLFUW[54]	UDUZOI/01[51,55]	0	0	NUHYEU[56]	0	1	0	1[b]
2-Methoxy	0	0	LANLUJ[57]	0	0	0	0	0	1	0
3-Methoxy	1	1	1	0	0	0	0	0	0	0
4-Methoxy	0	0	KIZYOJ/01[46,58]	0	0	0	0	0	1	0[b]
2-Methyl	1	1	0	1	1	1	0	1	1	0[b]
3-Methyl	1	1	1	0	0	0	0	0	0	0[b]
4-Methyl	1	1	OFOKOK[59]	0	0	0	0	0	0	1[b]

a - Reported crystal structure of a salt; b – Results from cocrystallisation screen performed by Stephen P. Stokes, UCC, unpublished work;

Table 2: Experimental results and literature results for cocrystal combinations in Quadrant 2 (1 indicates formation of a new solid form, 0 indicates no evidence of change observed in PXRD or IR).

Compound Name / Acid R-Group	Caffeine	Capryllic Acid	L-Glutamic Acid	Nicotinic Acid	Pyridoxine	3,3'-Thiodipropionic Acid	Malic Acid	Trans-Aconitic Acid	L-Ascorbic Acid	Riboflavin
2-Nitro	1	0	0	1	0	0	0	0	0	0
3-Nitro	1	0	1	0	0	0	0	1	0	0
4-Nitro	0	0	1	0	VUKPAU[a,60]	0	0	0	0	0
2-Hydroxy	1	0	0	0	0	0	0	0	0	0
3-Hydroxy	1	0	1	2	0	1	1	1	0	0
4-Hydroxy	1	0	0	0	VUKPEY[a,60]	0	0	0	0	0
2-Fluoro	1	0	1	1	1	1	0	0	0	0
3-Fluoro	1	0	0	0	1	0	0	0	0	0
4-Fluoro	0	0	1	0	0	0	0	0	0	1
2-Amino	1	0	0	1	0	0	0	0	0	0
3-Amino	0	0	0	0	0	0	0	0	0	0
4-Amino	0	0	0	SESLIM[61]	VUKPIC[a,60]	0	1	1	0	0
2-Methoxy	1	0	1	0	0	0	0	0	0	0
3-Methoxy	1	0	0	0	0	0	0	0	0	0
4-Methoxy	0	0	0	0	0	0	0	0	0	0
2-Methyl	1	0	1	1	1	1	1	1	1	0
3-Methyl	1	0	0	0	0	0	0	0	0	0
4-Methyl	0	0	0	0	0	0	0	0	0	0

a - Reported crystal structure of a salt;

Table 3: Experimental results and literature results for cocrystal combinations in Quadrant 3 (1 indicates formation of a new solid form, 0 indicates no evidence of change observed in PXRD or IR).

Compound Name / Amide R-Group	Nicotinamide	4,4'-Bipyridyl	Fumaric Acid	Salicylic Acid	Urea	Benzamide	Oxalic Acid Dihydrate	Hydro-quinone	2-Pyrrolid-inone
2-Nitro	0	1	0	0	0	0	1	0	2
3-Nitro	1	1	1	1	1	2	2	1	1
4-Nitro	1	1	0	1	0	1	1	0	1
2-Hydroxy	1	1	2	EGIQAO[62]	2	0	1	0	1
4-Hydroxy	1	1	GESBEN[63] GESCIS[63]	1	1	0	GERZUA[63]	2	1
3-Fluoro	0	1	1	0	0	0	1	1	1
4-Fluoro	0	1	1	1	0	0	1	0	1
2-Amino	0	1	1	1	0	0	1	0	1
3-Amino	0	1	1	2	0	0	1	1	2
4-Amino	1	1	1	0	1	0	1	0	1
2-Methoxy	0	1	1	1	1	0	1	1	2
3-Methoxy	0	0	1	0	0	0	1	0	1
4-Methoxy	0	1	1	1	0	0	1	1	1
2-Methyl	0	0	0	0	0	1	1	0	1
3-Methyl	0	1	1	1	0	2	1	1	1
4-Methyl	0	0	1	1	0	1	GENLUH[64]	1	1

Table 4: Experimental results and literature results for cocrystal combinations in Quadrant 4 (1 indicates formation of a new solid form, 0 indicates no evidence of change observed in PXRD or IR).

Compound Name / Acid R-Group	Caffeine	Capryllic Acid	L-Glutamic Acid	Nicotinic Acid	Pyridoxine	3,3'-Thiodipropionic Acid	Malic Acid	Trans-Aconitic Acid	L-Ascorbic Acid	Riboflavin
2-Nitro	0	0	0	0	0	0	0	0	0	0
3-Nitro	1	0	0	1	0	0	1	0	1	0
4-Nitro	1	0	0	0	0	1	1	0	0	0
2-Hydroxy	1	0	0	0	0	0	0	0	0	0
4-Hydroxy	1	0	0	1	0	0	1	0	0	0
3-Fluoro	1	0	0	0	0	0	1	1	0	0
4-Fluoro	0	0	0	0	0	1	1	1	0	0
2-Amino	1	0	0	0	0	1	1	0	0	0
3-Amino	0	0	0	0	0	0	0	1	1	0
4-Amino	0	0	0	0	0	0	1	0	0	0
2-Methoxy	0	0	0	0	0	0	0	1	0	0
3-Methoxy	0	0	0	0	0	1	1	0	0	0
4-Methoxy	1	1	1	1	1	1	1	1	1	0
2-Methyl	0	0	0	0	0	0	0	1	0	0
3-Methyl	0	0	0	0	0	0	0	0	0	0
4-Methyl	1	0	0	1	0	0	1	0	1	0

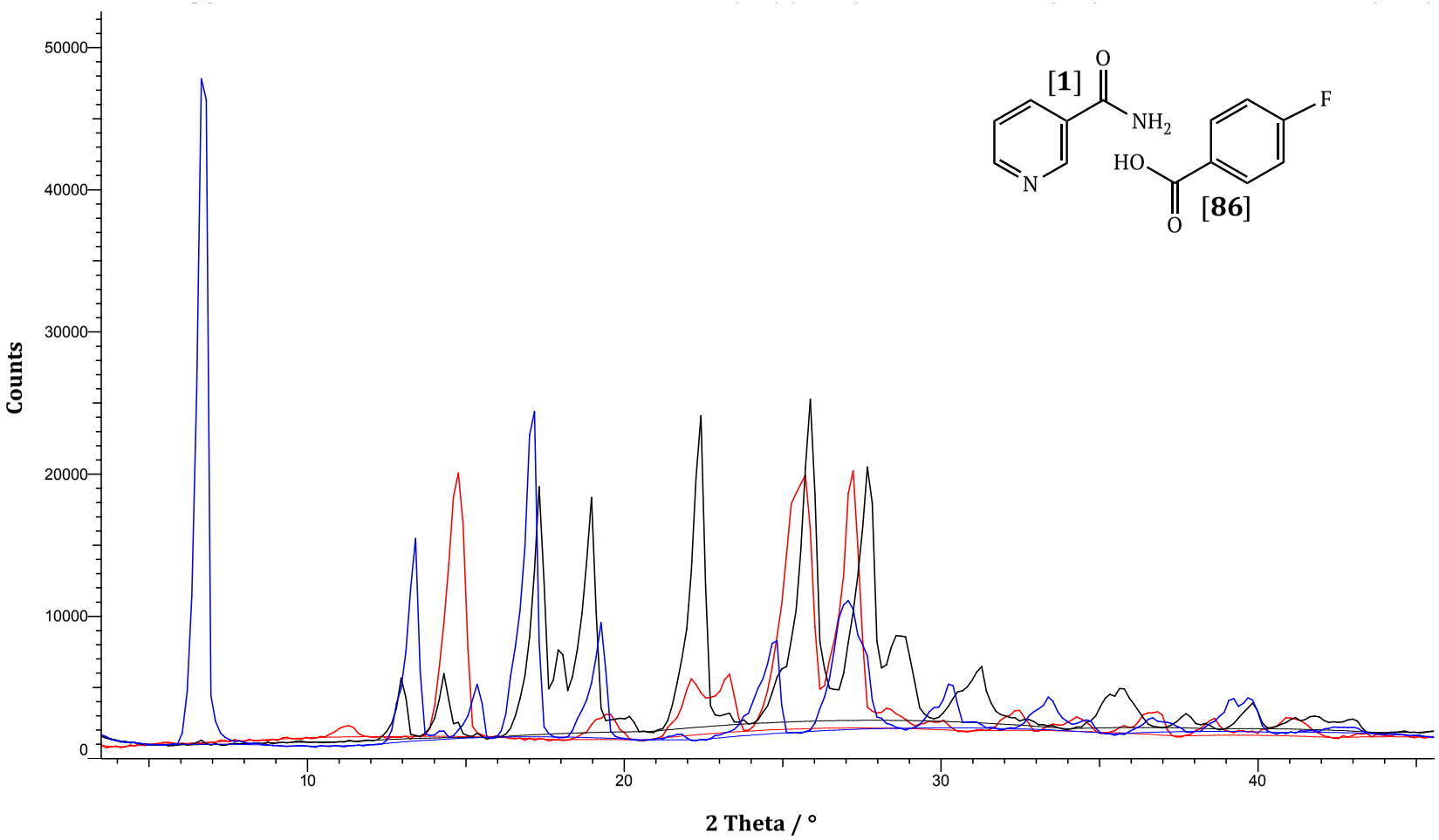


Figure 5: PXRD overlay of neat grinding product [1.86] (black), nicotinamide [1] (red), and 4-fluorobenzoic acid [86] (blue).

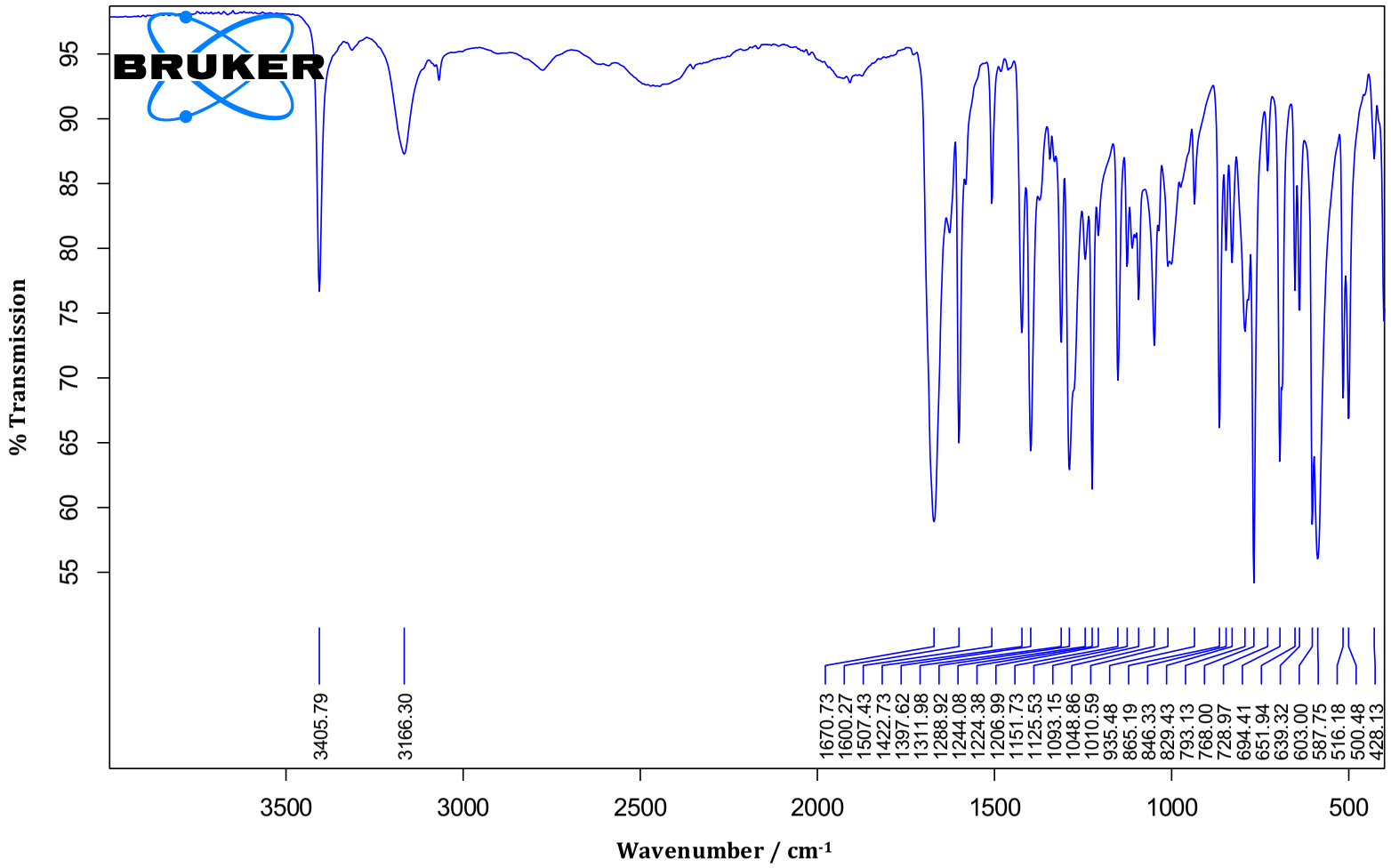


Figure 6: IR spectrum for neat grinding product 1.86.

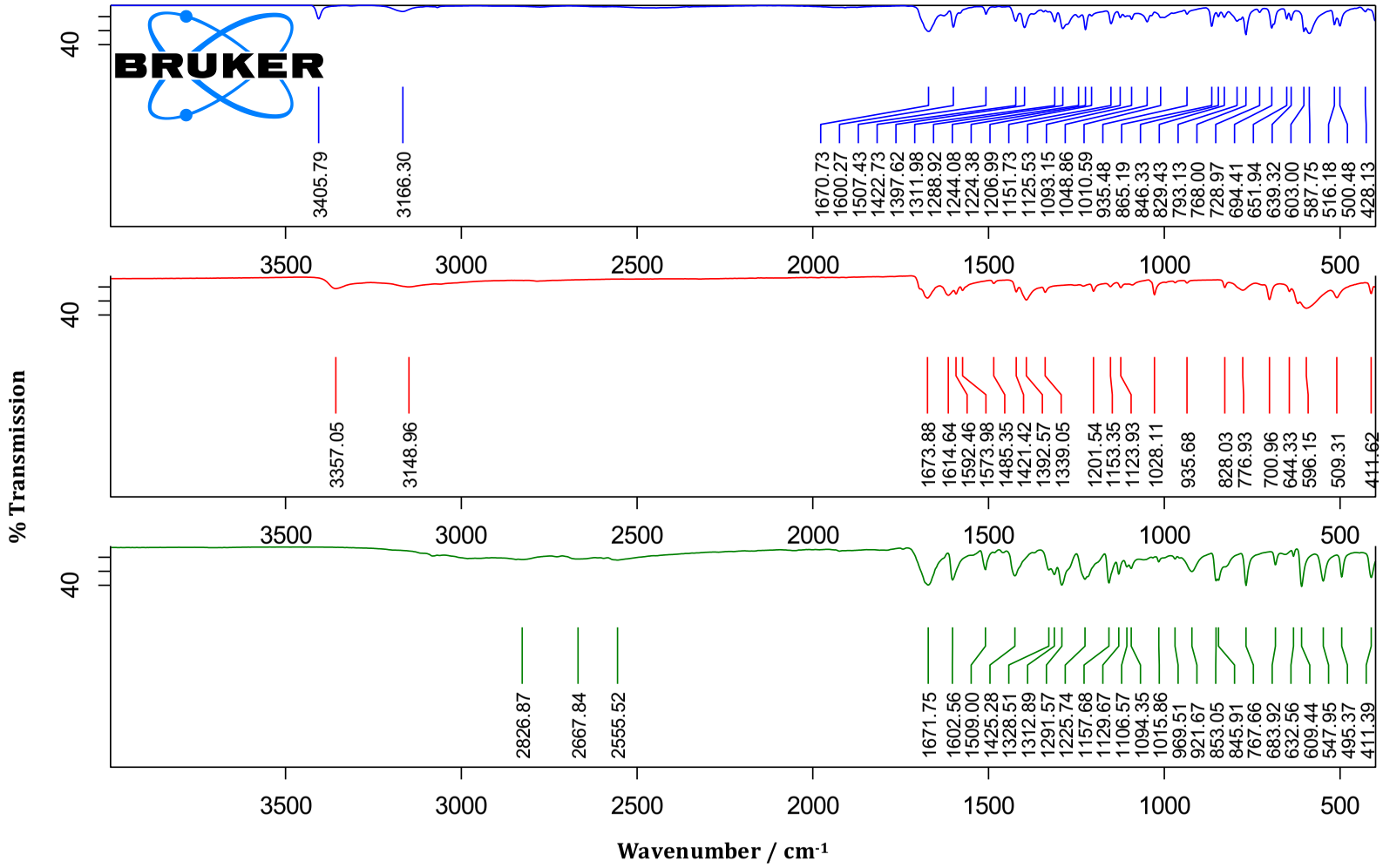


Figure 7: IR comparison of neat grinding product [1.86] (blue), nicotinamide [1] (red), and 4-fluorobenzoic acid [86] (green).

4.4 Design and Implementation of an SVM Classifier for Cocrystal Prediction

4.4.1 Preparation of Molecular Descriptors

The SVM classifier was prepared at the University of Oxford, using the Python coding language, the RDKit toolkit[8] was used to create a library of 191 standard molecular descriptors for each molecule in the data matrix. During development, it was observed that certain functional groups, namely tertiary-amines, amides and ethers, were incorrectly categorised in the code. The descriptors for these functional groups were corrected to accurately identify the functionalities, and novel descriptors for aryl-halide functional groups were also designed. This resulted in a library of 195 descriptors for each molecule.

In order to tailor the programme to interpret the formation of a cocrystal, it was decided to incorporate a library of α_2^H and β_2^H hydrogen bond propensity values from Hunter's table (Section 1.1.2),[15] which were subsequently used as a cocrystal descriptor for a pairwise relationship of hydrogen bond interactions between the selected functional groups. Descriptors for each cocrystal pair were created *via* concatenation of the individual descriptors for the cocrystal components, creating a library of 391 descriptors for each potential cocrystal.

4.4.2 SVM Classification

The SciKit learn package[7] was used for the supervised training of the SVM with the binary data matrix. A common method that can be used to train the SVM with a data matrix such as this is to withhold a portion of the data from training (commonly a row or column), and to use the trained model to predict the outcomes for this small subset of the data.

In order to minimise the level of bias in the study, an external set of data was used as the validation set for prediction. The study by Wood *et al.*[10] which incorporated paracetamol [**131**] with 35 different coformers (13 successful) was used for this purpose.

4.4.3 SVM Accuracy

The cross validation accuracy of the model for prediction on the small subset of data was promising at 75.0 ± 1.4%. However, the predictive accuracy on the paracetamol [**131**] validation set was less so, at 64.7%. A more powerful application of this methodology was to 'enrich' the number of successful hits in a ranked list of the potential coformers.

The ranked list was created according to the probability of cocrystal formation generated by the SVM classifier. In this list, of the 13 successful cocrystals in the validation set, 9 were placed in the top 11 of the ranked list (Figure 8). From this ranked list, the enrichment factor (EF_{25}) of the model could be calculated, which describes the number of correctly classified data points in the top 25% of the data. The enrichment factor was determined to be 2.6 over random classification, therefore, use of this algorithm gives a 2.6-fold increase in identification of successful cocrystals over a random selection from the coformers.

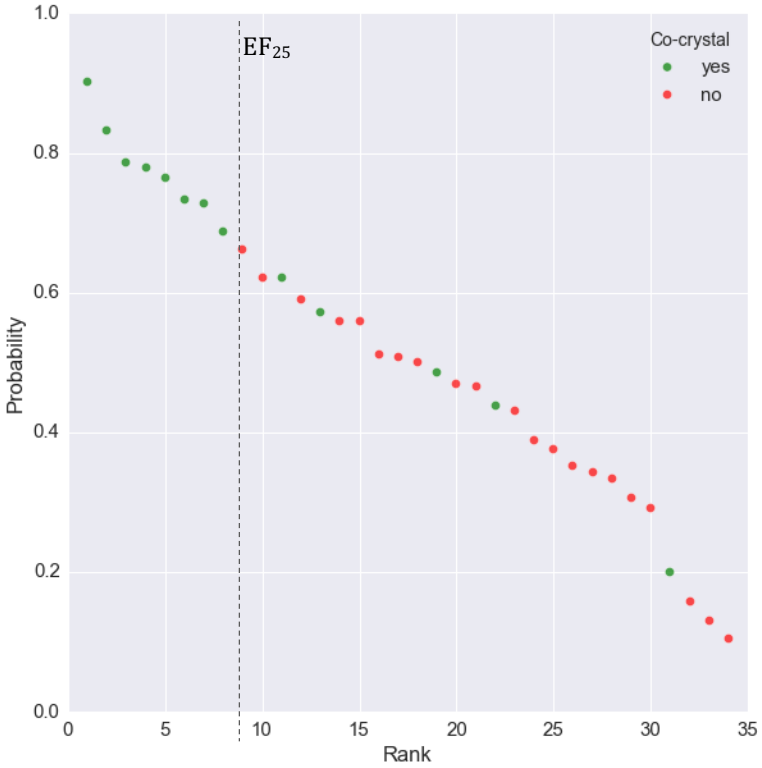


Figure 8: Enriched list of coformers by probability ranking of the paracetamol [131] validation set. Green dots indicate successful cocrystals, red dots indicate unsuccessful cocrystals, EF_{25} shown by the dashed line.

The receiver operating characteristic (ROC) curve was also calculated (Figure 9), which illustrates the performance of the model in the binary classification

problem. The area under this curve (AUC) provides an indication of the probability that the model will rank a randomly chosen positive result higher than a randomly chosen negative result. This value was determined to be 0.86, which compared favourably with the 0.66 value determined for the HBP methods used by Wood *et al*. in their report.

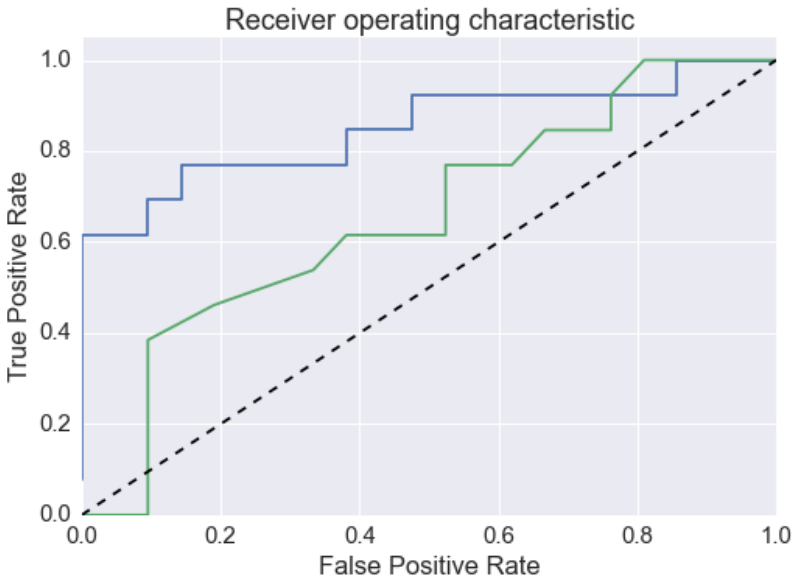


Figure 9: Receiver operating characteristic (ROC) curve for this work (blue) and the HBP method used by Wood *et al*.[10] (green), random classification is shown by dashed line.

The reduced level of accuracy in the validation test set on **131** was attributed to a high number of false positive values in the confusion matrix (unsuccessful cocrystals that were predicted to form) [Table 5]. The report by Wood *et al*.[10] has described false positive values in a study such as this as 'less of a concern' since one cannot guarantee that a cocrystal will never form. The low number of false negative values in the confusion matrix is very encouraging (Table 5).

Table 5: Confusion matrix for SVM model on paracetamol [131] test set.

	Cocrystal predicted to form	**Cocrystal predicted *not* to form**
Co-crystal formed	12 True Positive	1 False Negative
No co-crystal formed	11 False Positive	10 True Negative

4.5 Conclusions

This work has demonstrated the potential application of SVM classification in developing a ranked list of coformers for a cocrystallization screen. The model performs favourably when compared to other computational methods such as HBP analysis, while providing the additional advantage of being computationally efficient. The use of such a tool in conjunction with experimental cocrystallization screening could significantly speed up the development of novel cocrystals while minimizing the number of unsuccessful outcomes.

4.6 Characterisation of Quadrant 1 Cocrystals

Having identified 260 novel materials in the data matrix for the SVM, efforts turned toward the characterization of some of these cocrystals. In order to create a manageable workload, only the novel materials identified in Q1 were targeted for full characterization. Q1 contained a total of 85 cocrystals, of which 35 were already contained in the literature, and a further 9 were unpublished results from our research group.[65] The remaining group of 41 materials will be discussed in this section.

4.6.1 Differential Scanning Calorimetry and Melting Points of Q1 Cocrystals

DSC was performed on each grinding product to determine the melting point of the materials. The 41 grinding products were dissolved in each of ethanol, methanol and acetonitrile (where soluble), and crystallized as per section 2.7.4. DSC was performed as per section 2.7.3.

For the majority of the materials, a strong singular endotherm indicating the melting point was observed, however, in a few cases the DSC pattern was complex, and so a visual melting point was determined as a supporting analytical technique. For example, in the case of **109.79**, urea and 2-methylbenzoic acid, there were two endotherms observed in the DSC pattern, a smaller one at approx. 85 °C, and a stronger endotherm at 102-104 °C (Figure 10). Visual melting point analysis of this compound confirmed that the melting point was the latter

endotherm, observed at 103-105 °C. Results of the DSC analysis of these materials is shown in Table 6 and Table 7.

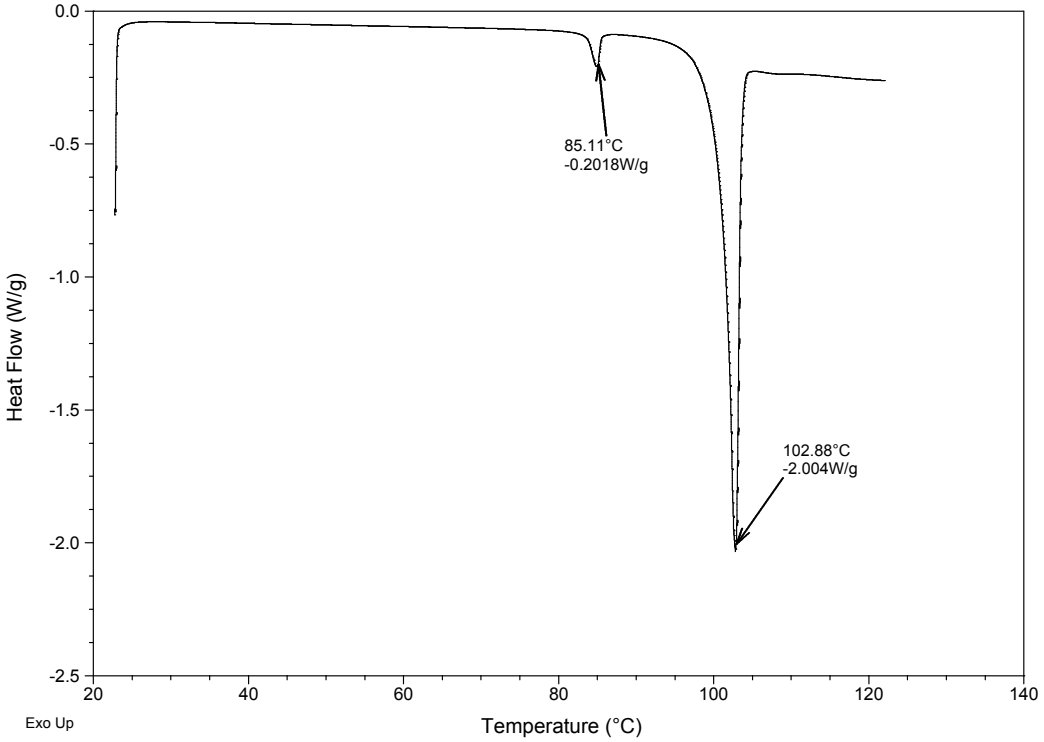


Figure 10: DSC pattern obtained for 109.79.

Intriguingly, the distribution of the melting points for cocrystals determined in this study when compared to the melting points of the individual coformers, are not in agreement with the statistical analysis of cocrystal melting points performed by Perlovich in 2015.[66] Perlovich analysed the melting points of 727 cocrystals from the CSD,[9] and determined that in the majority of cases (55.3%) the melting point of the resultant cocrystal is between that of the individual coformers.

In our study, it has been observed that the melting point of the cocrystal is between that of the coformers in just 22.0% of cases (although the number of data points in Perlovich's work was higher).[66] The majority of the cocrystals determined in this study display melting points lower than either of the coformers (65.8%), and just 12.2% display melting points higher than the components (Figure 11).

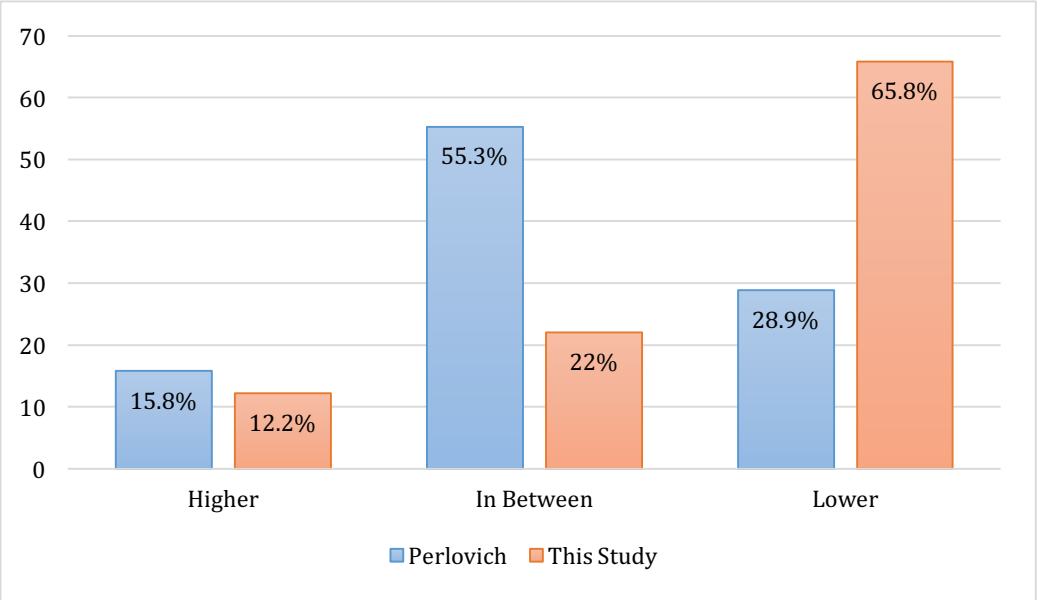


Figure 11: Comparison of cocrystal melting point distributions between this work (orange) and that determined by Perlovich (blue).[66]

If melting points are considered as a reflection of the overall lattice energies of these materials, it is interesting to observe that, in this instance, the majority of the cocrystals would be interpreted to be less stable than both of their coformers. This is a striking correlation when one considers that the rationale behind cocrystal prediction used by Karamertzanis *et al.* in 2009 was to assess the predicted lattice energies of the cocrystal by comparison to the coformers, with more stable lattices interpreted as cocrystals that were predicted to form.[12]

The observations for Q1 cocrystals are consistent with the melting points determined for the salsalate cocrystals in Chapter 2, of which two thirds displayed melting points lower than either of the coformers. The completed set of DSC endotherms (and melting points, where applicable) for materials studied in Q1 is shown in Table 6 and Table 7.

Table 6: DSC endotherms for novel solid forms identified in Q1.

Compound number	DSC (endotherm)	m.p. of coformer 1	m.p. of coformer 2
1.121	87 °C, 98 °C, 109-114 °C visual mp: 90-93 °C	Nicotinamide 129 °C[67]	2-Nitrobenzoic Acid 143-145 °C[68]
1.104	171.5-172.5 °C		3-Nitrobenzoic Acid 141.5 °C[69]
1.84	93-96 °C		2-Fluorobenzoic Acid 120 °C[70]
1.86	141-142 °C		4-Fluorobenzoic Acid 182 °C[70]
1.77	110-114 °C		2-Aminobenzoic Acid 145 °C[71]
1.124	97-99 °C		3-Methoxybenzoic Acid 106.5 °C[72]
1.79	78-80 °C		2-Methylbenzoic Acid 103-105 °C[73]
1.80	97 °C, 102 °C, 107-109 °C visual mp: 122-124 °C		3-Methylbenzoic Acid 111 °C[74]
1.81	83-85 °C		4-Methylbenzoic Acid 175 °C[75]
61.121	126-129 °C	Isonicotinamide 159 °C[76]	2-Nitrobenzoic Acid 143-145 °C[68]
61.77	113-115.5 °C		2-Aminobenzoic Acid 145 °C[71]
61.78	107-109 °C		3-Aminobenzoic Acid 174 °C[77]
61.124	130-133 °C		3-Methoxybenzoic Acid 106.5 °C[72]
61.79	107-110 °C		2-Methylbenzoic Acid 103-105 °C[73]
61.80	135-136 °C		3-Methylbenzoic Acid 111 °C[74]
61.81	181-183 °C, 185-187 °C visual mp: 182-184 °C		4-Methylbenzoic Acid 175 °C[75]
15.84	88-90 °C, 101-103 °C 101-103 °C	4,4'-Bipyridyl 112 °C[78]	2-Fluorobenzoic Acid 182 °C[70]
15.124	122-124 °C		3-Methoxybenzoic Acid 106.5 °C[72]
15.80	92-94 °C		3-Methylbenzoic Acid 111 °C[74]
108.82	194-196 °C	Fumaric Acid 284 °C[79]	3-Hydroxybenzoic Acid 199 °C[80]
108.78	172-178 °C visual mp: 177-178 °C		3-Aminobenzoic Acid 174 °C[77]
108.79	145-149 °C		2-Methylbenzoic Acid 103-105 °C[73]

Table 7: DSC endotherms for novel solid forms identified in Q1.

Compound number	DSC (endotherm)	m.p. of coformer 1	m.p. of coformer 2
60.78	128-130 °C	Salicylic Acid 159 °C[81]	3-Aminobenzoic Acid 174 °C[77]
60.79	92.5 °C, 94-96 °C visual mp: 93-97 °C		2-Methylbenzoic Acid 103-105 °C[73]
109.121	63-67 °C	Urea 133 °C[82]	2-Nitrobenzoic Acid 143-145 °C[68]
109.122	149-152 °C		3-Nitrobenzoic Acid 141.5 °C[69]
109.84	83-85 °C		2-Fluorobenzoic Acid 182 °C[70]
109.79	85 °C, 102-104 °C visual mp: 103-105 °C		2-Methylbenzoic Acid 103-105 °C[73]
13.121	87.5-90 °C	Benzamide 128 °C[75]	2-Nitrobenzoic Acid 143-145 °C[68]
13.84	70-73 °C		2-Fluorobenzoic Acid 120 °C[70]
13.85	80-82 °C		3-Fluorobenzoic Acid 124 °C[70]
13.86	107 °C, 110-111 °C visual mp: 112-114 °C		4-Fluorobenzoic Acid 182 °C[70]
110.82	161-163 °C	Oxalic Acid Dihydrate 100-101 °C[83]	3-Hydroxybenzoic Acid 199 °C[80]
110.77	visual mp: 149-151 °C		2-Aminobenzoic Acid 145 °C[71]
110.2	visual mp: > 160 °C decomposition		4-Aminobenzoic Acid 186 °C[84]
110.79	visual mp: 96-99 °C		2-Methylbenzoic Acid 103-105 °C[73]
5.82	152-155 °C	Hydroquinone 170.5 °C[85]	3-Hydroxybenzoic Acid 199 °C[80]
5.123	visual mp: 99-100 °C		2-Methoxybenzoic Acid 101.5 °C[86]
5.105	visual mp: 97-99 °C		3-Methoxybenzoic Acid 106.5 °C[72]
5.79	101-103 °C		2-Methylbenzoic Acid 103-105 °C[73]
5.81	99-101 °C		4-Methylbenzoic Acid 175 °C[75]

4.6.2 Single Crystal Analysis

Single crystal analysis was performed to characterise the crystal structures of as many of the materials in Q1 as possible. Of the 41 materials in the study, 13 were fully characterised using SCXRD, and a further 2 were characterised using this method, but generated structures that had significant levels of disorder (**1.81** and **1.84**).

4.6.2.1 Cocrystals of Nicotinamide [1]

Nicotinamide [**1**] (3-pyridinecarboxamide), also known as niacinamide, is the water soluble, amide metabolite of nicotinic acid [**91**], (Niacin, Vitamin B3).[87] Nicotinamide is used in dermatology for the treatment of acne vulgaris and rosacea,[88] as a lightening agent and antimicrobial,[89] and has been shown to have photo-protectant properties in the battle against skin cancer.[87] It is commonly used as a coformer in cocrystallisation studies, and has been incorporated into several pharmaceutically important solid forms.[90,91]

Studies have determined that Form I is the more stable of the two polymorphic forms of **1**, however, the $R_2^2(8)$ dimer is only observed in Form II.[92] The more stable Form I contains an infinite C(4) chain extending along the *c*-axis using one of the two amide N-H, and the other N-H involved in an acid-pyridyl interaction forming a spiral down the *a*-axis (Figure 12).

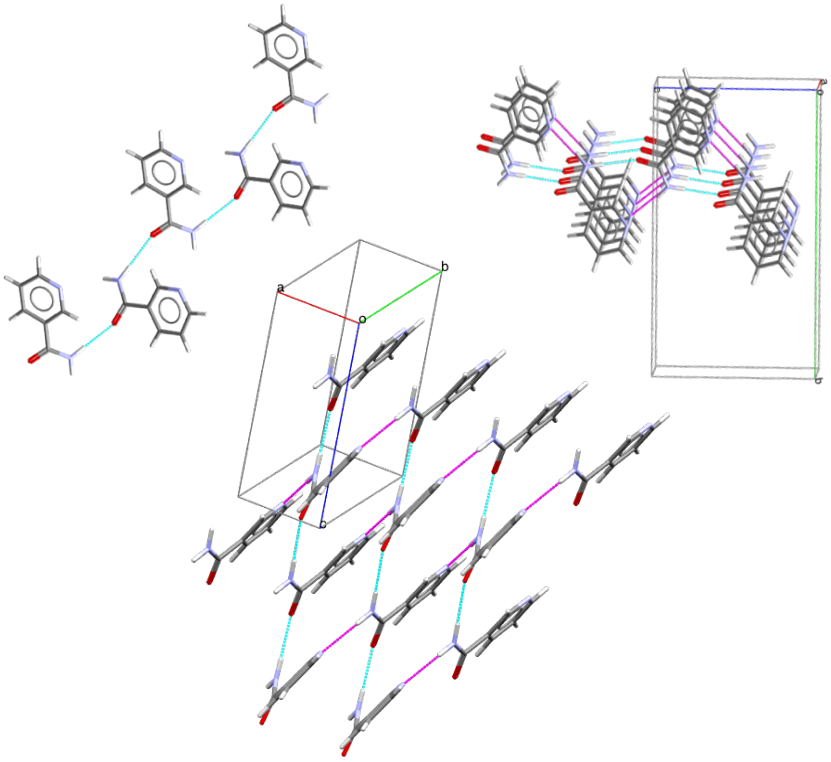


Figure 12: Hydrogen bonding motifs in nicotinamide [1] Form I (NICOAM01).[93]

Form II displays the $R_2^2(8)$ motif characteristic to the amide functional group, with 4 independent molecules in the asymmetric unit. The dimer units are bridged *via* N-H···O=C and N-H···N=C interactions, forming a binary level $R_4^4(20)$ ring motif (Figure 13).

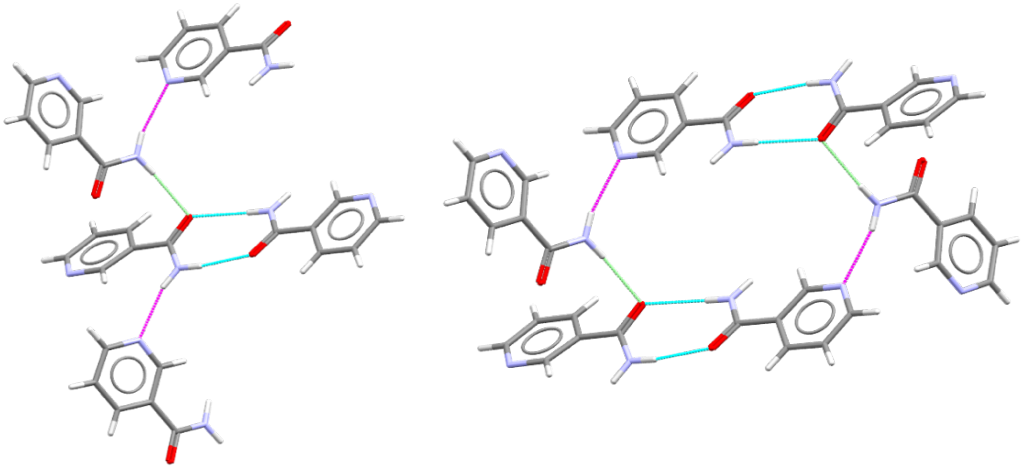


Figure 13: Hydrogen bond interactions observed in 1, Form II (NICOAM04).[94]

There are 127 crystal structures published in the CSD[9] including **1**, of which 118 contain more than one unique molecule (cocrystals), and 22 are salts. 85 of the nicotinamide cocrystals have a carboxylic acid functional group in the coformer,

similar to those determined in this work. 29 cocrystals centre around an $R_2^2(8)$ heteromeric dimer formed between **1** and the carboxylic acid (frequency 35%, Figure 14 [a]), and 78 of these form the dependable acid-pyridyl interaction (frequency 95% Figure 14 [b]). 36 structures contain $R_2^2(8)$ homodimer [frequency 44%] (Figure 14 [c]). The distribution of these interactions combinatorially is shown in Table 8.

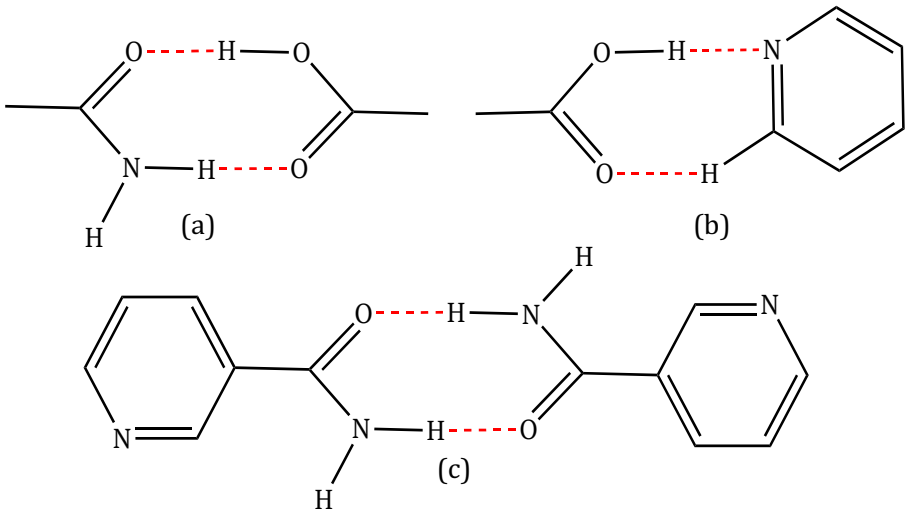


Figure 14: Supramolecular synthons observed in cocrystals of 1.

Table 8: Combinations of primary interactions observed in nicotinamide carboxylic acid cocrystals.

Motif 1	Motif 2	No. of Cocrystals
$R_2^2(8)$ Homodimer	$R_2^2(8)$ Heterodimer	1[a]
Acid-Pyridyl	$R_2^2(8)$ Heterodimer	25
$R_2^2(8)$ Homodimer	Acid-Pyridyl	36

a - This structure contains all three motifs;

Of the 9 new solid forms of nicotinamide identified in Q1 of this study, 6 were successfully characterised using SCXRD. Successful results for nicotinamide in Q1 are summarised in Table 9.

Table 9: Cocrystallisation results in Q1 for 1 (novel solid forms identified in this study highlighted in green).

	Coformer	Space Group	CSD[9] Refcode
121	2-Nitrobenzoic Acid[b]	$C2/c$	N/A
104	4-Nitrobenzoic Acid[b]	$P\text{-}1$	N/A
60	2-Hydroxybenzoic Acid	$P2_1/n$	SODDOF/01/02[30–32]
82	3-Hydroxybenzoic Acid	$P2_1/c$	XAQQIQ[33]
83	4-Hydroxybenzoic Acid	$C2/c$	RUYHEZ/01[43,44]
84	2-Fluorobenzoic Acid[b,c]	$P2_1/c$	N/A
86	4-Fluorobenzoic Acid[b]	$P2_1/c$	N/A
77	2-Aminobenzoic Acid[b]	$P2_1$	N/A
2	4-Aminobenzoic Acid[a]	$P\text{-}1$	ABULIU[53]
124	3-Methoxybenzoic Acid	-	N/A
79	2-Methylbenzoic Acid	-	N/A
80	3-Methylbenzoic Acid[b]	$P2_1/n$	N/A
81	4-Methylbenzoic Acid	-	N/A

a - Hydrated structure; b – Crystal structure determined using SCXRD;
c - Crystal structure was of poor quality

5 of the 6 new materials that were characterised in this work were neutral cocrystals, with the remaining structure a partial salt incorporating both neutral and ionised forms of the carboxylic acid. The cocrystal structure of **1.84** was confirmed via SCXRD, however, the structure was poor with a large level of rotational disorder around the phenyl-carboxyl bond and so was not included in this thesis (unit cell data is available in experimental Section 4.8).

The first two materials that will be discussed here display a common central motif of the homomeric nicotinamide $R_2^2(8)$ dimer, surrounded by additional secondary interactions. This parent dimer retention has been observed previously in cocrystals of structurally-related isonicotinamide [**61**].[40] This dimer is interesting, in that it structurally related to the less stable polymorph of nicotinamide, and also represents an energetically weaker interaction than the heteromeric $R_2^2(8)$ dimer could potentially form with the carboxylic acid coformer.[15]

The cocrystal of **1** with 4-nitrobenzoic acid [**1.104**] crystallizes in a 1:1 stoichiometric ratio in the triclinic space group $P\text{-}1$ [a = 7.1167(5) Å, b = 7.5590(5) Å , c = 12.8081(9) Å, α = 85.164(2)°, β = 75.933(2)°, γ = 85.895(2)°, V = 665.04(8) Å^3]. The central $R_2^2(8)$ dimer is capped by N-H···O=C interactions to

the carboxylic acids. The pyridyl moiety also accepts a hydrogen bond from the acid O-H creating a separate binary ring motif, an $R_4^4(20)$ tetramer (Figure 15).

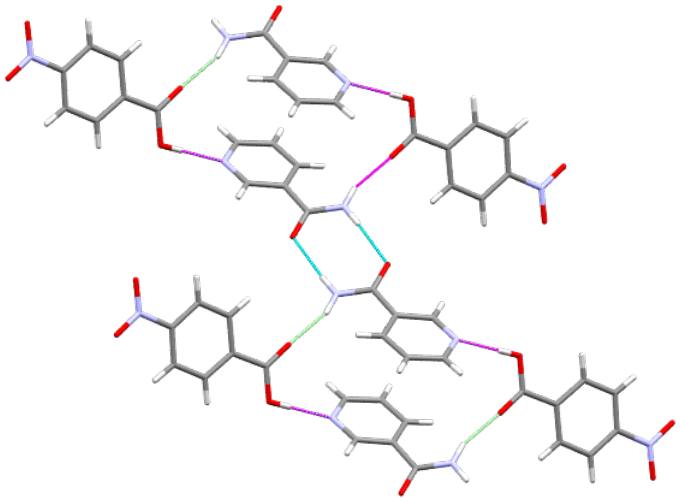


Figure 15: $R_4^4(20)$ tetramer formed in in the 1:1 cocrystal 1.104 [bonds shown in magenta, cyan and green].

In the 1:1 cocrystal of 1 with 4-fluorobenzoic acid [**1.86**], this central motif is also capped by similar carboxylic acid interactions, with further interactions observed to the pyridyl moiety also (Figure 16).

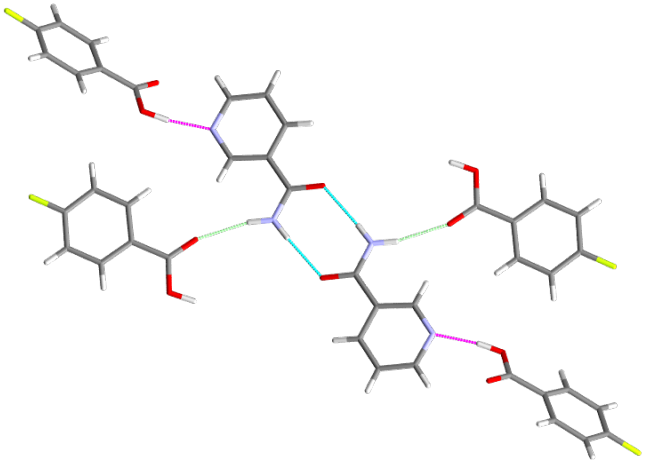


Figure 16: Central motifs displayed in 1:1 cocrystal 1.86 [bonds shown in magenta, cyan and green].

Further to this central motif observed in both materials, the nitro groups in **1.104** (4-nitrobenzoic acid) participate in two additional interactions here, at an approximate interaction distance of 2.6 Å in both cases. Both interactions occur to aromatic C-H atoms, one to nicotinamide and the other to the 4-nitrobenzoic acid (Figure 17, blue and orange). The latter interaction is complemented by a

C-H···O-H interaction to the acid hydroxyl, forming an $R_2^2(10)$ dimer interaction overall (2.6 Å, Figure 17, orange).

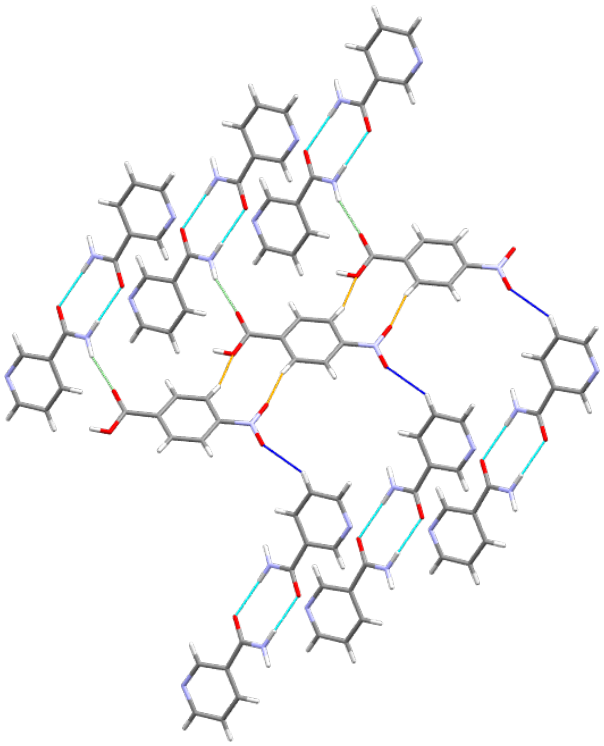


Figure 17: Nitro to C-H interactions in 1:1 cocrystal 1.104 [hydrogen bonds in orange and blue].

In **1.86** (4-fluorobenzoic acid) there are halogen bonds (2.4 Å) observed as secondary motifs in the structure, forming a binary $R_4^4(22)$ tetramer in conjunction with a C-H···O=C bond (2.6 Å) from the fluorophenyl ring to the carbonyl on the amide (Figure 18).

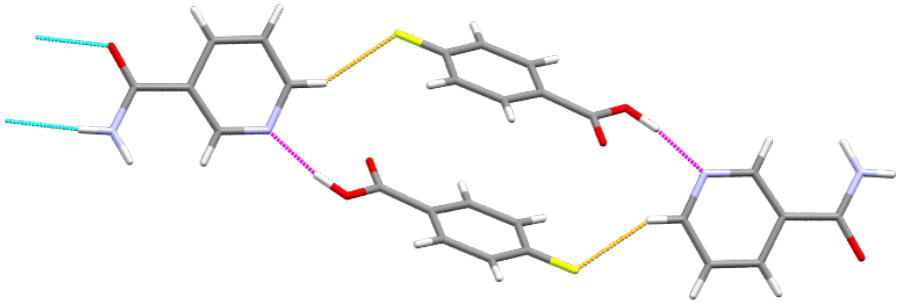


Figure 18: Binary tetramer formed in 1:1 cocrystal 1.86 [bonds shown in magenta, and orange].

At an architectural level, the interactions observed in **1.86** result in spiralling of the molecules along the *c*-axis (Figure 19), creating a complex criss-crossed pattern overall.

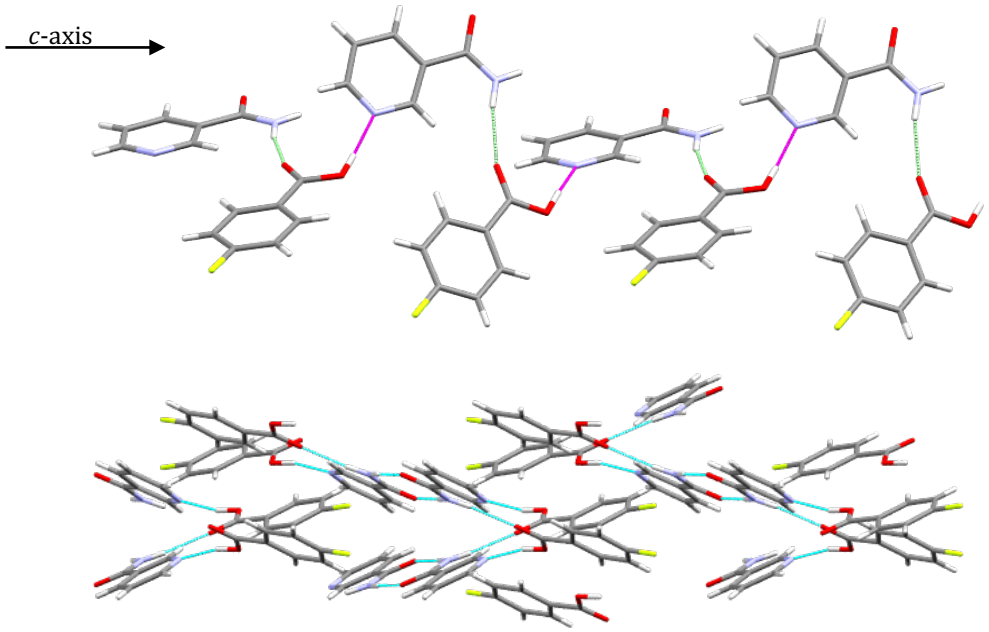


Figure 19: Spiralling of molecules along the *c*-axis in 1:1 cocrystal 1.86 (top) and overall packing motifs viewed down the *c*-axis (bottom) [bonds shown in magenta, cyan and green].

The cocrystal of **1** with 3-methylbenzoic acid [**80**] was solved in monoclinic $P2_1/c$ [a = 13.629(16) Å, b =7.151(8) Å , c = 13.651(15) Å, β = 115.649(17)°, V = 1199.0(2) Å^3]. This material contains the component molecules in a 1:2 ratio, and breakage of the strong $R_2^2(8)$ homomeric dimer occurs in favour of the more energetically-favoured $R_2^2(8)$ heteromeric acid-amide dimer (Figure 20). Secondary interaction occurs from the acid hydroxyl to the pyridyl nitrogen forming a spiral down the *b*-axis (Figure 21).

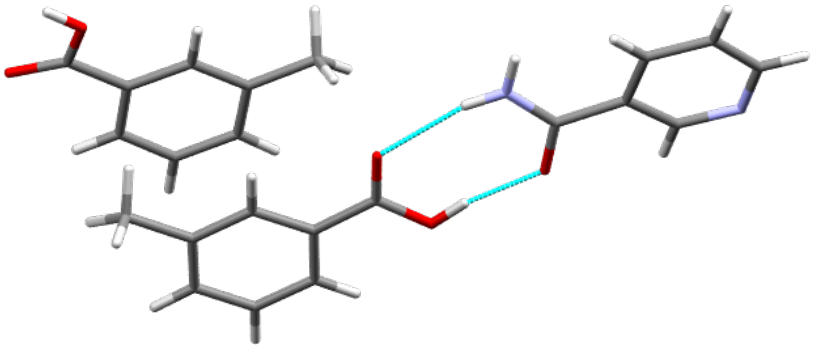


Figure 20: The asymmetric unit of the 1:2 cocrystal 1.80.

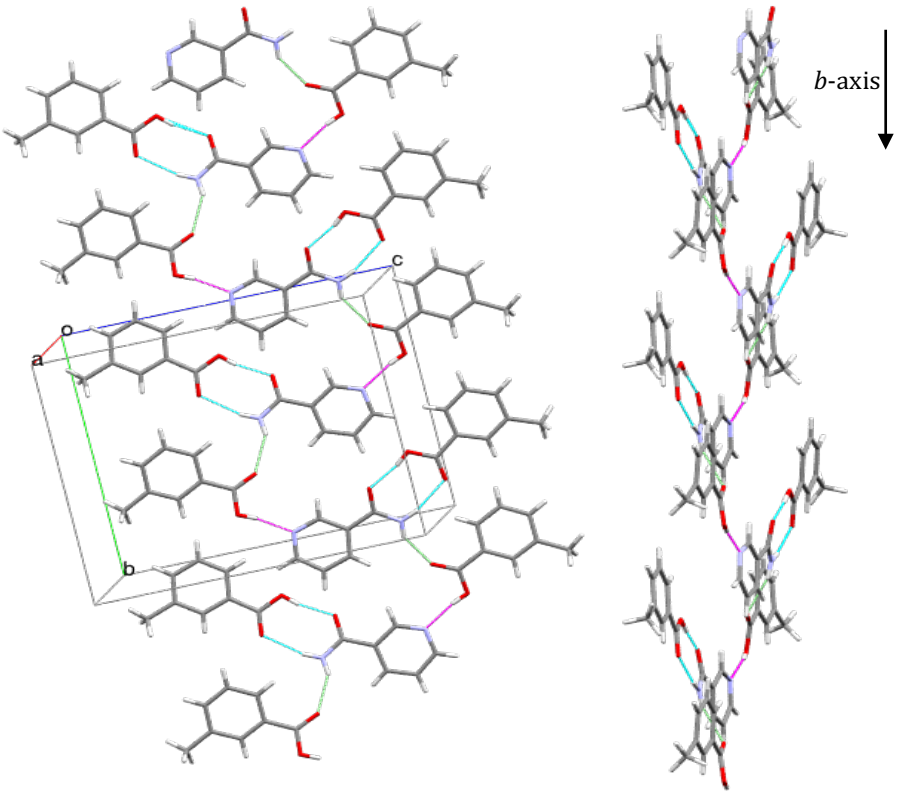


Figure 21: Spiralling of molecules down the *b*-axis in 1:1 cocrystal 1.80 [bonds shown in magenta, cyan and green].

Interestingly, the heteromeric dimer is not observed in the case of **1** with 2-aminobenzoic acid [**77**]; the structure has a 1:1 ratio and crystallizes in the monoclinic space group $P2_1$ [a = 10.479(2) Å, b =4.9873(9) Å , c = 12.644(3) Å, β = 109.361(5)°, V = 623.4(2) Å^3]. The strongest motifs present here are a C(4) chain formed between the molecules of **1**, reminiscent of the bonding observed in Form I, which are crosslinked *via* discrete bonds from the amide to the carbonyl of the carboxylic acid, and from the acid hydroxyl back to the pyridine nitrogen (Figure 22 and Figure 23). This combination of interactions creates a distinct spiralling interaction down the *b*-axis via the 2_1 screw axes (Figure 24).

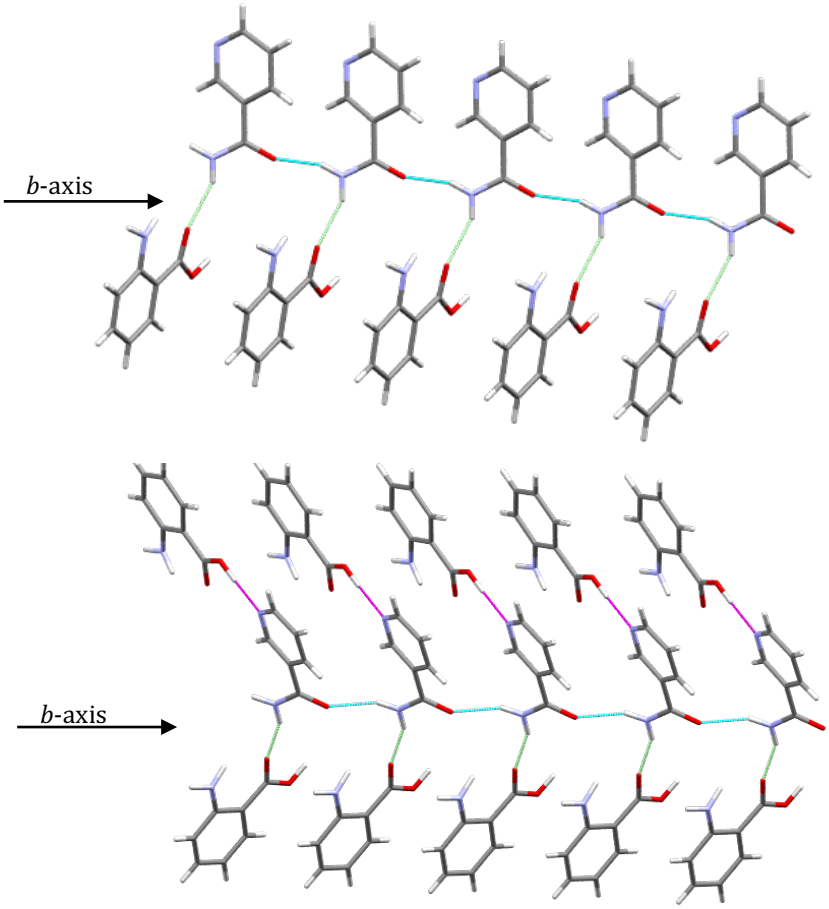


Figure 22: Crosslinked C(4) chain formed between nicotinamide molecules in 1:1 cocrystal 1.77.

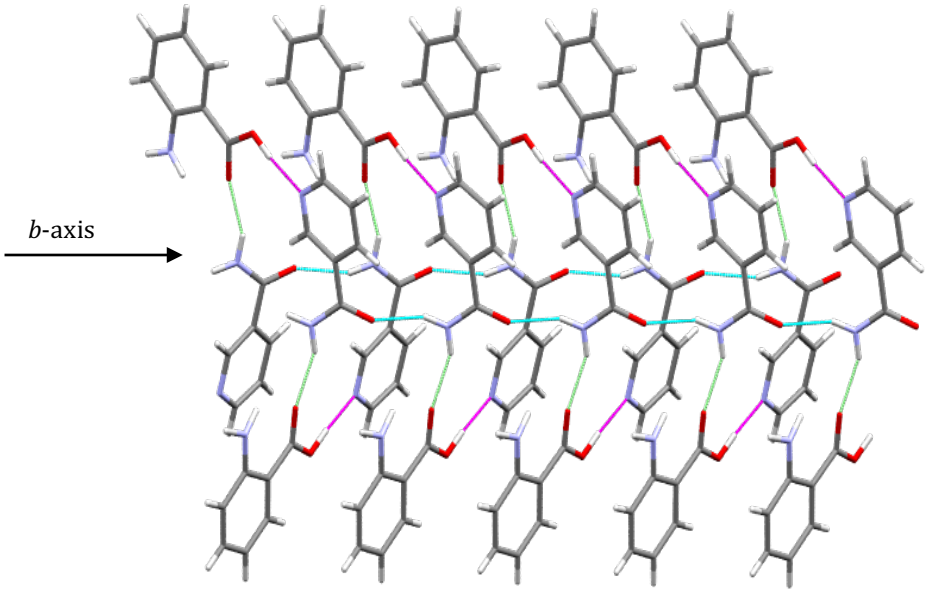


Figure 23: Spiralling of the molecules down the *b*-axis formed by crosslinked C(4) chains in 1:1 cocrystal of 1.77.

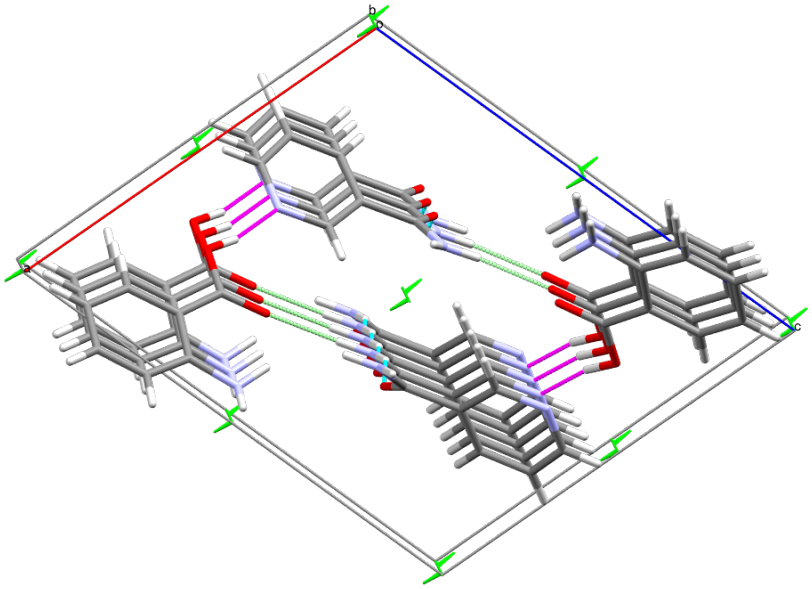


Figure 24: Spiralling of molecules down the *b*-axis in 1.77 (view down *b*-axis), [bonds shown in magenta, cyan and green, 2_1 screw axis in green].

This spiralling motif is complemented by an N-H···N amino-amino interaction. The amino hydrogen points toward the nearest amino nitrogen at a distance of 2.5 Å and an angle of 86° (Figure 25) creating a chain down the *b*-axis above and below the spiral shown in Figure 24.

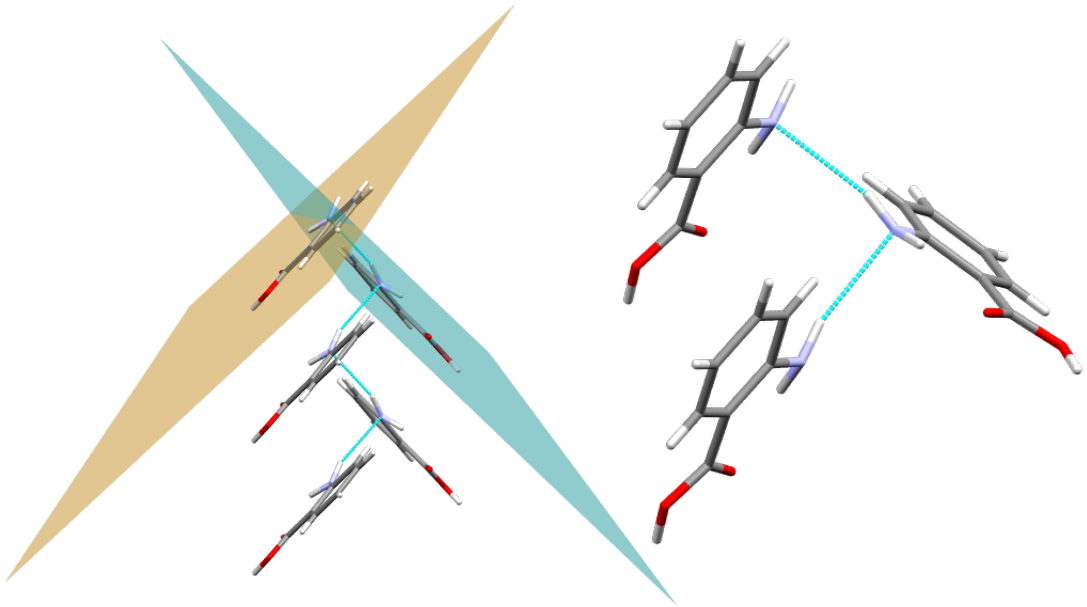


Figure 25: Perpendicular interactions between amino groups in 1.77.

An interesting structure determined in this series is the partial-salt structure of **1** with 2-nitrobenzoic acid [**121**]. This structure is in a 1:2 stoichiometric ratio, with one neutral and one anionic molecule of 2-nitrobenzoic acid, where the carboxylic acid proton has transferred to the pyridine nitrogen on **1**. This

material [**1.121**] crystallises in the monoclinic space group *C2/c* [a = 27.715(3), b = 7.0371(7), c = 21.947(2), β = 105.132(4), V = 4131.98 Å^3]. The nicotinamide molecules have formed the strong $R_2^2(8)$ dimer motif at the centre of the structure. Around this dimer, there are discrete interactions from the carboxylate group to the protonated nitrogen (Figure 26, magenta), from the neutral carbonyl to the amide N-H (Figure 26, green), and from the nitro group on the neutral carboxylic acid to the hydrogen *ortho* to the pyridine nitrogen (Figure 26, orange). Finally, the second carboxylate oxygen accepts a bond from the neutral hydroxyl hydrogen (Figure 27, orange).

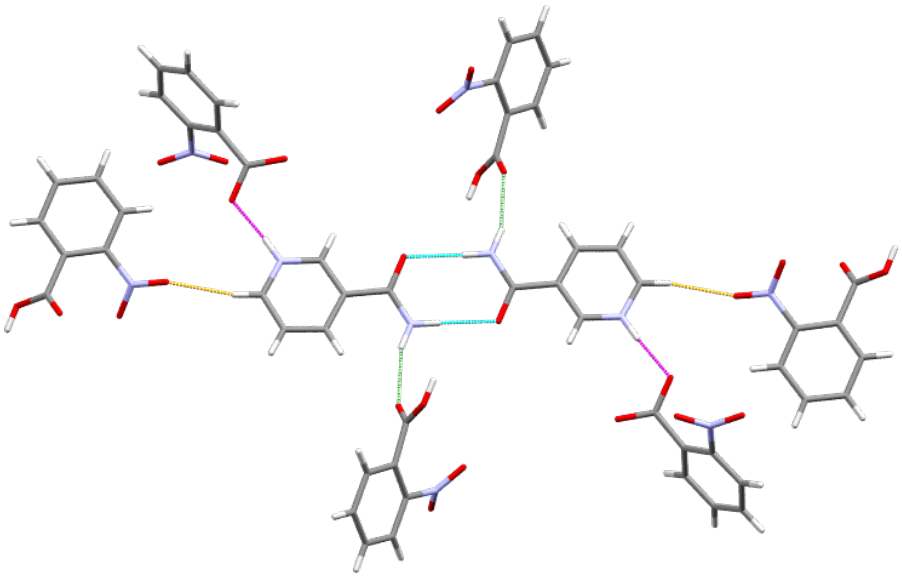


Figure 26: Interactions around the central dimer in 1:2 cocrystal 1.121 [bonds shown in magenta, cyan, orange and green].

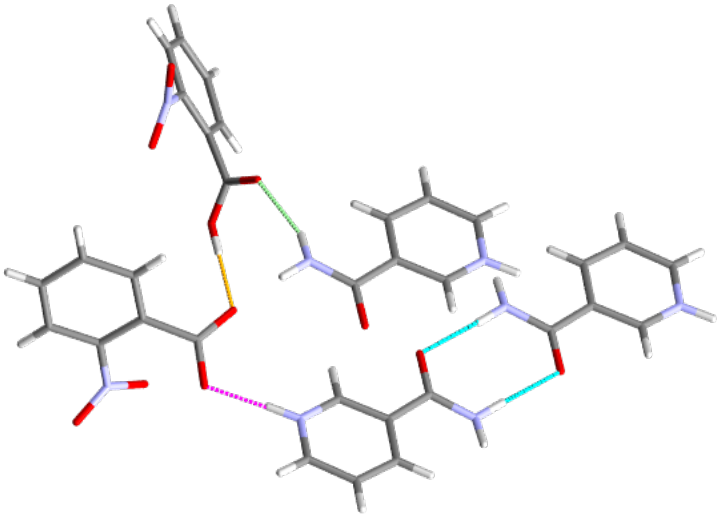


Figure 27: Interactions observed in 1:2 cocrystal of 1.121 [bonds shown in magenta, cyan, orange and green].

The overall crystal packing results in the nicotinamide dimers forming down the central line and capped at the sides by the various interactions to the carboxylic acid/carboxylate molecules (Figure 28 and Figure 29).

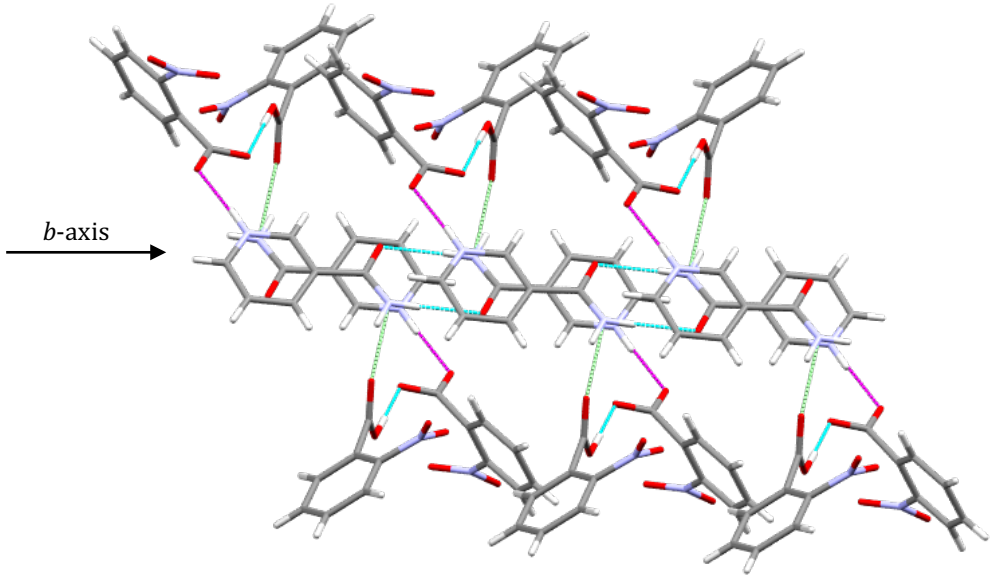


Figure 28: Nicotinamide dimers forming a central pattern in the packing of 1:2 cocrystal of 1.121 [bonds shown in magenta, cyan, and green].

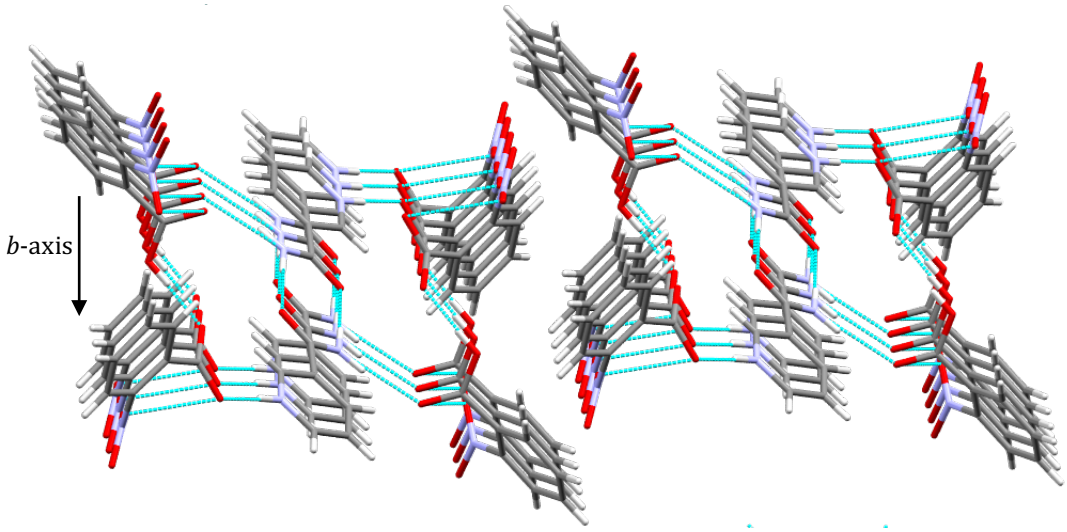


Figure 29: Alternative view of central dimers along *b*-axis in 1.121.

There is one other partial salt structure of **1** contained in the CSD.[9] PARPUV is the crystal structure of pyridinium-3-carboxamide nicotinamide chloride monohydrate. The structure contains a central dimer formed between one cationic and one neutral molecule of nicotinamide [**1**], which is capped by interactions to the anionic chloride and water molecules at the centre of the unit cell. There is also a pyridinium-pyridine interaction between the neutral and anionic nicotinamide molecules (Figure 30).

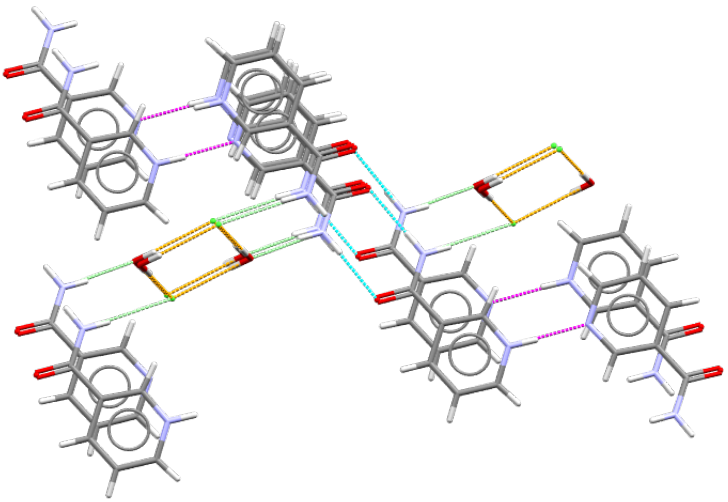


Figure 30: Hydrogen bonding interactions observed in PARPUV.

Direct comparison of these two crystal structures is difficult as the neutral and ionised materials in this structure are distributed differently. However, it is interesting to observe that in the presence of strong ionic interactions, the $R_2^2(8)$ homomeric dimer prevails as the primary hydrogen bonding interaction in this case also.

4.6.2.2 Cocrystals of Isonicotinamide [61]

Isonicotinamide (4-pyridine-carboxamide, **61**) is the structural analogue of nicotinamide [**1**]. It is a reliable, commonly used coformer in cocrystallisation studies.[40,95–98] Isonicotinamide has proven to be highly polymorphic, with 5 forms published in the literature.[25,94,99] Form I contains an $R_2^2(8)$ dimer and a C(4) chain, but these motifs do not form the amide ladder (as discussed for amides Section 2.3.2), instead forming an extended 3-dimensional array due to alternating orientation of the dimers (Figure 31). The remaining forms (II-V) all form large 2-dimensional layers of $R_4^4(18)$ tetramers as formed by amide C(4) chains and amide-pyridyl interactions (Figure 32).

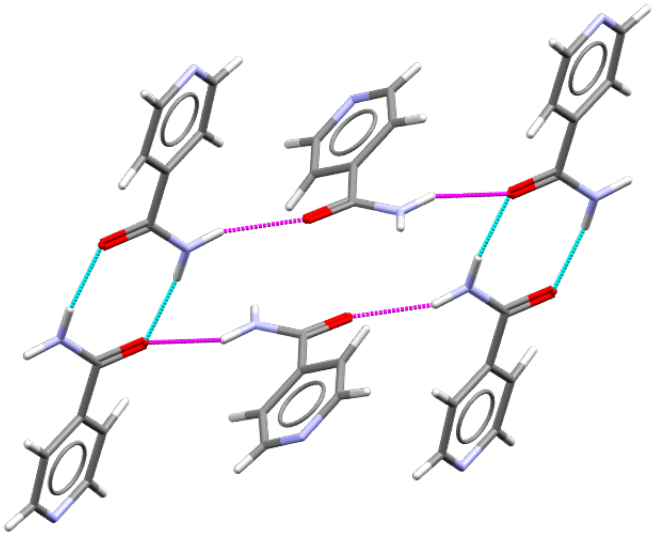


Figure 31: Interactions observed in isonicotinamide [61] Form I [EHOWIH01].[25]

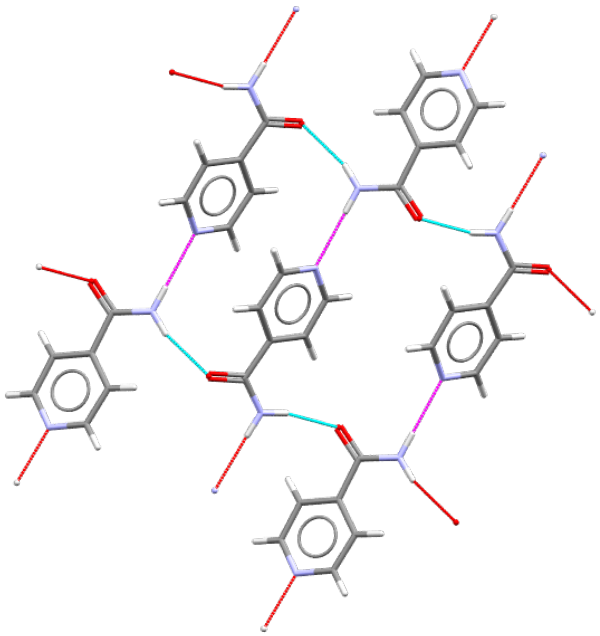


Figure 32: Extended 2-D array of tetramers in 61 Form II-V (Form V shown) [EHOWIH05].[99]

Retention of the Form I homodimer in cocrystallisation is a consistent feature in many cocrystals involving carboxylic acids. There are 114 cocrystals in the current CSD[9] featuring **61** with an acid coformer, of which 85 retain the isonicotinamide $R_2^2(8)$ dimer and form heteromeric interactions *via* other means. Aakeröy and coworkers[40] described two predictable motifs in cocrystals of **61** with a variety of carboxylic acid coformers; (a) the heteromeric carboxylic acid-pyridine hydrogen bond and (b) an $R_2^2(8)$ homodimer of **61**.[40] In 2012,

Desiraju and Tothadi noted that the latter interaction occurs commonly in cocrystals of **61** with carboxylic acids.[98]

Isonicotinamide proved to be the most successful coformer utilised in this study, with a total of 16 successful results from the 18 combinations in Q1 alone. 9 materials were already contained in the literature, and 7 novel solid forms were identified using the combination of PXRD and IR analysis. 3 were fully characterised using SCXRD (**61.121**, **61.77** and **61.124**), 1 other structure was confirmed as a 1:1 cocrystal using SCXRD [**61.81**] however, the structure was of poor quality and was not included in this work, despite multiple attempts at SCXRD. Results for **61** in Q1 are summarised in Table 10.

Table 10: Cocrystallisation results in Q1 for 61 (novel solid forms identified in this study highlighted in green).

	Coformer	Space Group	CSD[9] Refcode
121	2-Nitrobenzoic Acid[a]	$P2_1/c$	N/A
122	3-Nitrobenzoic Acid	$I2/a$	ASAXOH[22]
104	4-Nitrobenzoic Acid	P-1	AJAKEB[25]
60	2-Hydroxybenzoic Acid	$P2_1/c$ P-1	XAQQEM[33] QAFTID[34]
82	3-Hydroxybenzoic Acid	$C2/c$	LUNMEM[40]
83	4-Hydroxybenzoic Acid	$P2_1/n$	VAKTOR[45]
84	2-Fluorobenzoic Acid	P-1	HANHEL[47]
85	3-Fluorobenzoic Acid	$C2/c$	CACGUK[47]
86	4-Fluorobenzoic Acid	$C2/c$	ASAXUN/01[22,47]
77	2-Aminobenzoic Acid[a]	$P2_1/c$	N/A
78	3-Aminobenzoic Acid	-	N/A
2	4-Aminobenzoic Acid	$P2_1/c$	SOLFUW[54]
124	3-Methoxybenzoic Acid[a]	P-1	N/A
79	2-Methylbenzoic Acid	-	N/A
80	3-Methylbenzoic Acid	-	N/A
81	4-Methylbenzoic Acid[a,b]	P-1	N/A

a – Crystal structure determined using SCXRD; b - Crystal structure was of poor quality.

The two most prevalent architectural features of isonicotinamide cocrystals in the literature are the homomeric $R_2^2(8)$ dimer (85 examples, frequency 74.6%), and the acid-pyridyl interaction (106 examples, frequency 97.2%) and to a lesser extent the acid-amide heteromeric dimer (22 examples, frequency 19.3%).

The nitro-substituted cocrystals (**61.121**, **61.122**, and **61.104**)[22,25] display a similar level of complexity in their bonding patterns to that observed in similar

cocrystals of nicotinamide (**1.121** and **1.104**).[25] It is interesting to observe the presence of the isonicotinamide $R_2^2(8)$ homomeric dimer even in the presence of nitro groups, this contrasts directly with the trend of dimer breakage observed for other aromatic amides (Section 2.3.2). In all three cases, the homomeric dimer is predictably capped by interactions to the carboxylic acid, and the free amide N-H bonds to the carboxyl oxygen in a hierarchical manner (Figure 33, Figure 34, and Figure 35).

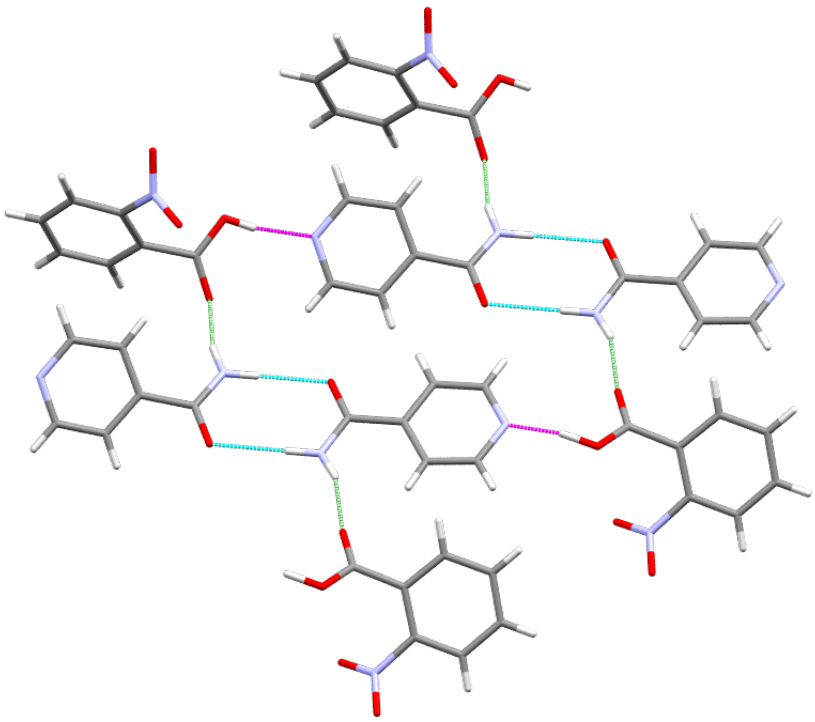


Figure 33: Primary hydrogen bonding motifs observed in 1:1 cocrystal 61.121 [bonds shown in magenta, cyan, and green].

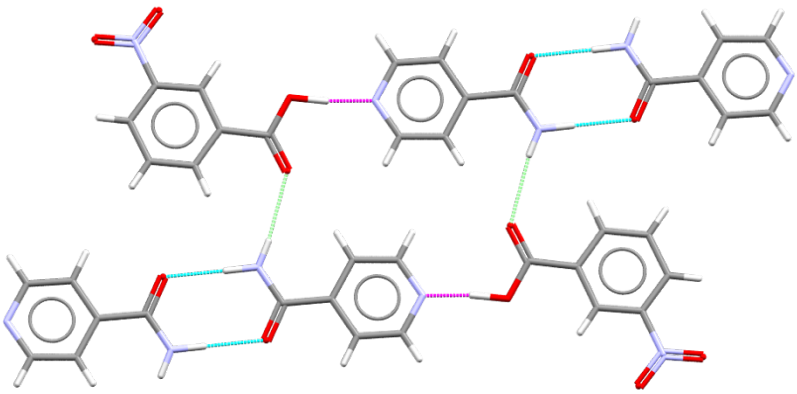


Figure 34: Primary hydrogen bonding motifs observed in 1:1 cocrystal 61.122 [ASAXOH][22] [bonds shown in magenta, cyan, and green].

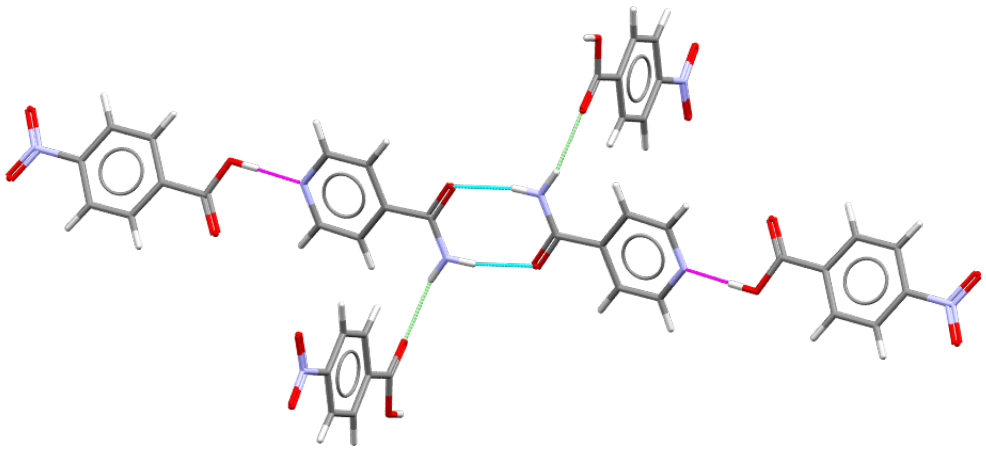


Figure 35: Primary hydrogen bonding motifs observed in 1:1 cocrystal 61.104 [AJAKEB][25] [bonds shown in magenta, cyan, and green].

Interestingly, there is no participation of the nitro groups in the hydrogen bonding of **61.121**, contrasting directly with that of **61.122** [ASAXOH][22] and **61.104** [AJAKEB],[25] both of which display C-H···O=N interactions, which combine at binary levels to form interesting ring structures (Figure 36 and Figure 37).

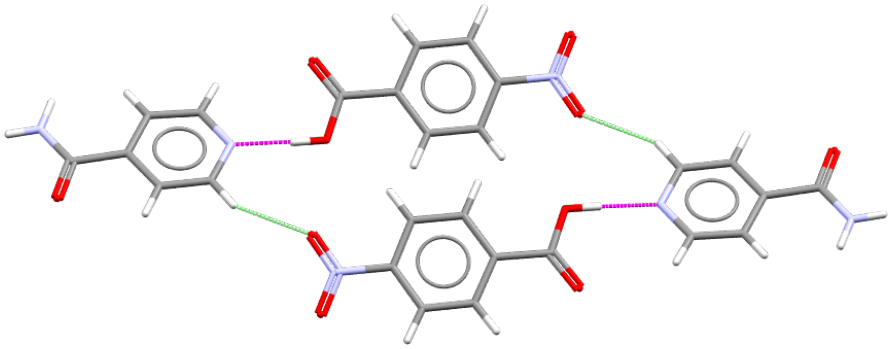


Figure 36: C-H···O=N interactions in 61.122 [ASAXOH][22] forming a binary $R_4^4(24)$ ring structure [hydrogen bonds in magenta and green].

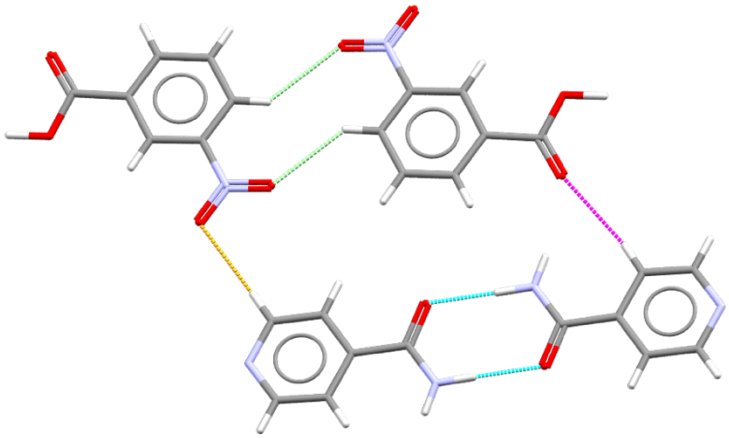


Figure 37: C-H···O=N interactions in 61.104 [AJAKEB][25] forming $R_2^2(10)$ dimer interaction (green), and discrete interaction (orange).

The second group of structures are those of isonicotinamide with amino substituted benzoic acids **61.77**, **61.78** and **61.2**.[54] Unfortunately, the structure of **61.78** could not be determined and so the observed motifs cannot be compared here. The motifs observed for **61.77** and **61.2** (SOLFUW)[54] display few similarities. In the case of **61.78**, the central isonicotinamide dimer is retained, the anticipated acid-pyridyl interaction is observed, and the amino group donates a hydrogen bond to the acid carbonyl (Figure 38).

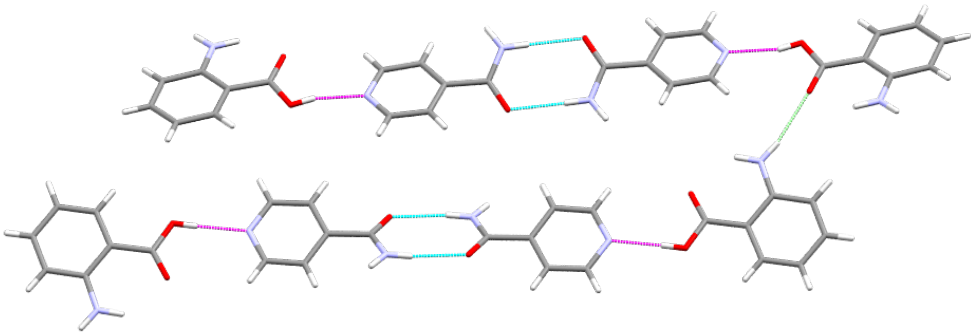


Figure 38: Hydrogen bonding interactions observed in 61.77 [isonicotinamide dimer (cyan), acid-pyridyl (magenta) and amino-carbonyl (green)].

In contrast, in the case of **61.2** (SOLFUW)[54] the dimer is broken in favour of the heteromeric acid-amide dimer, which would be predicted traditionally with this combination of molecules. This material crystallizes in a 2:1 ratio of acid to amide, also displaying a homomeric acid-acid dimer. Secondly, the highly persistent acid-pyridyl interaction does not occur in this material, with an amino-pyridyl interaction in its place (Figure 39).

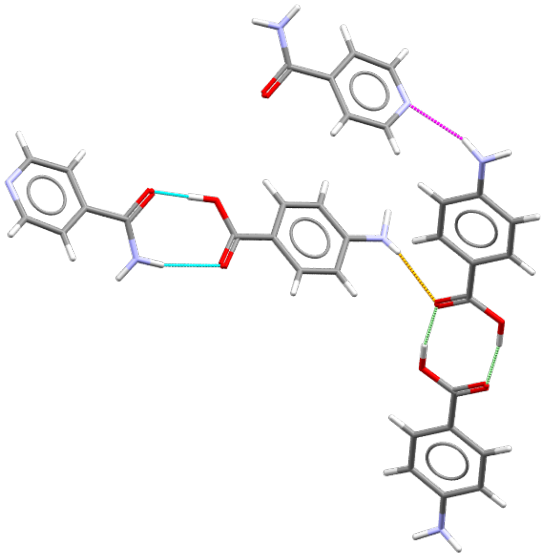


Figure 39: Hydrogen bonding interactions in 61.2 (SOLFUW)[54] [acid-amide dimer (cyan), amino-pyridyl (magenta), acid-acid dimer (green), amino-carbonyl (orange)].

The methoxy-substituted benzoic acid series yielded one successful result, a 1:2 cocrystal of isonicotinamide with 3-methoxybenzoic acid [**61.124**]. The observed motifs displayed in this material are consistent with that which would be traditionally predicted, the acid groups of two separate molecules of **124** participate individually in both the $R_2^2(8)$ acid-amide dimer, and the acid pyridyl interactions, thereby conferring the 1:2 stoichiometry observed here (Figure 40). Apart from these interactions, there is a moderate (2.5 Å) C-H···O-C hydrogen bonding interaction from one methoxy group on **124**, acting as donor, to another molecule of **124**, acting as acceptor (Figure 41).

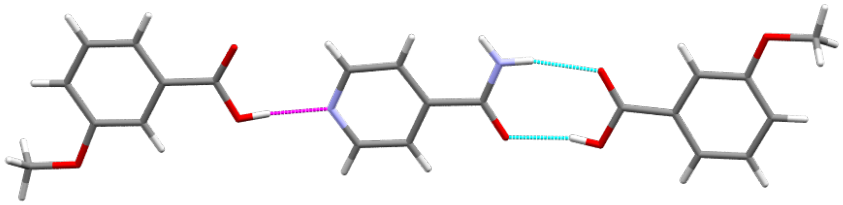


Figure 40: Acid-amide dimer (cyan) and acid-pyridyl (magenta) interactions observed in 61.124.

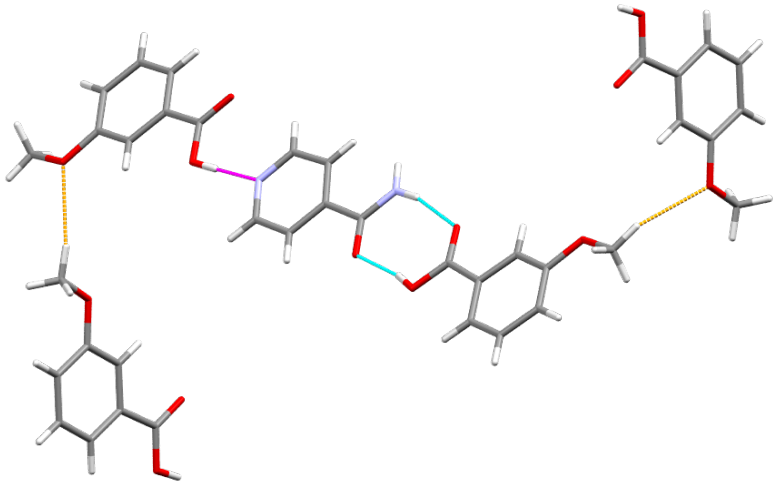


Figure 41: Methoxy C-H···O-C interactions observed in cocrystal 61.124.

4.6.2.3 Cocrystals of 4,4'-Bipyridyl [15]

4,4'-Bipyridyl [**15**] is a particularly useful coformer, commonly observed in cocrystals in the literature (517 multicomponent materials in the CSD).[9] This material proved highly successful as a coformer in this study also, forming cocrystals with 15 of the 18 benzoic acids, 12 of which were already available in the literature (Table 11).

Table 11: Cocrystallisation results in Q1 for 15 (novel solid forms identified in this study highlighted in green).

	Coformer	Space Group	CSD[9] Refcode
121	2-Nitrobenzoic Acid	$P2_1/c$	GAWJEU[21]
122	3-Nitrobenzoic Acid	$C2/c$	PUJHUY[23]
104	4-Nitrobenzoic Acid	$P2_1/n$	DAQZIF/01[26,27]
60	2-Hydroxybenzoic Acid	$P2_1/c$ P-1	KONZEU[35] KOPKEH[35]
82	3-Hydroxybenzoic Acid	P-1	HONVAI/01[41,42]
83	4-Hydroxybenzoic Acid	$P2_1/n$ $C2/c$ $P1$	EPUPUB[46] EPUPUB01[46] EPUQEM[46]
84	2-Fluorobenzoic Acid[a]	$P2_1/n$	N/A
77	2-Aminobenzoic Acid	$P2_1/c$	NINXOZ[48]
78	3-Aminobenzoic Acid	P-1 P-1	LEGPIY[50] UDUZIC[51]
2	4-Aminobenzoic Acid	$P2_1/n$	UDUZOI/01[51,55]
123	2-Methoxybenzoic Acid	$P2_1/c$	LANLUJ[57]
124	3-Methoxybenzoic Acid	-	N/A
105	4-Methoxybenzoic Acid	$P2_1/c$ $P2_1/c$	KIZYOJ/01[46,58]
80	3-Methylbenzoic Acid	-	N/A
81	4-Methylbenzoic Acid	$P2_1/c$	OFOKOK[59]

a – Crystal structure determined using SCXRD;

Due to the reduced number of available hydrogen bond acceptor sites on **15**, the landscape of hydrogen bonding interactions within this set of materials is strikingly simple. Of the 20 structures in the series (polymorphs included), all but two structures display **15** lying at a symmetrical position at the centre of the structure, with capping of **15** by acid-pyridyl interactions (Figure 42). This motif displays a 95% frequency in this series, and had been used successfully to synthesise cocrystals of salsalate, as discussed earlier (Chapter 3).

Secondly, there are several examples of mixed interactions to the opposing sides of **15**, where the availability of hydrogen bond donors on the coformer competes with the acid group for the interaction. For example, the 1:2 cocrystal with 3-hydroxybenzoic acid (**15.82**, HONVAI01)[42] displays hydrogen bonds from both the acid and the hydroxyl moieties to the pyridyl nitrogen, creating infinite chains (Figure 43).

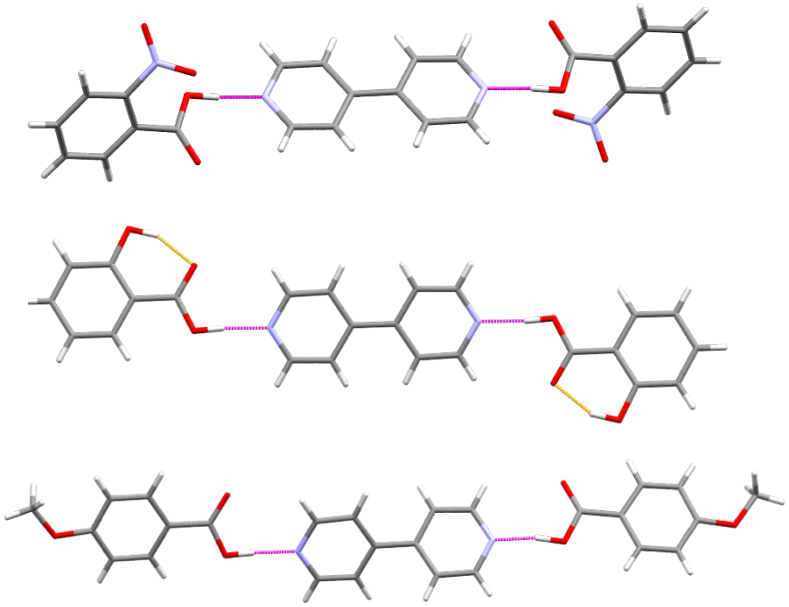


Figure 42: Persistence of the acid-pyridyl motif in 15 cocrystals GAWJEU [15.121, top], KONZEU [15.60, centre] and KIZYOJ [15.105 bottom].[21,35,58]

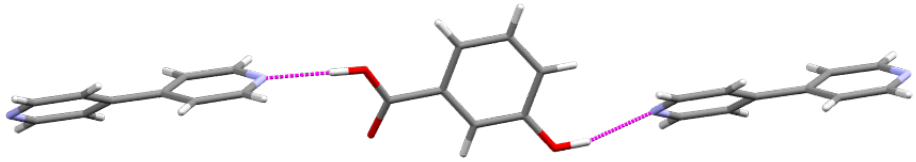


Figure 43: Infinite chains formed in 15.82 (hydrogen bonds in magenta) [HONVAI01].[42]

The amino group can also disrupt the acid-pyridyl interaction, for example in the case of 3-aminobenzoic acid [**15.78**] one of the pyridyl nitrogens accepts a hydrogen bond from the amino group, with the second amino N-H participating in an $R_2^2(14)$ ring motif at either side (Figure 44).

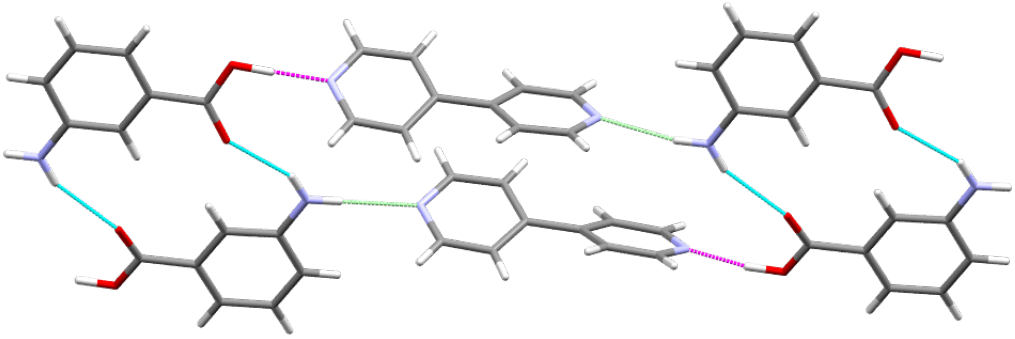


Figure 44: Hydrogen bonding interactions observed in the 1.5:1 cocrystal 15.78 (LEGPIY).[50]

Interestingly, the novel cocrystal characterised in this study (**15.84**, 2-fluorobenzoic acid) does not contain a competitor for the acid-pyridyl interaction, and yet displays an interaction to one side of **15** only, not the capping motif most commonly observed with these systems. The other side of the molecule is involved in a weak (2.8 Å) C(3) chain interaction, creating a spiral down the *b*-axis (Figure 45). This set of hydrogen bonding motifs is striking as it is in contrast to expectations from Etter's rules and Hunter's table.

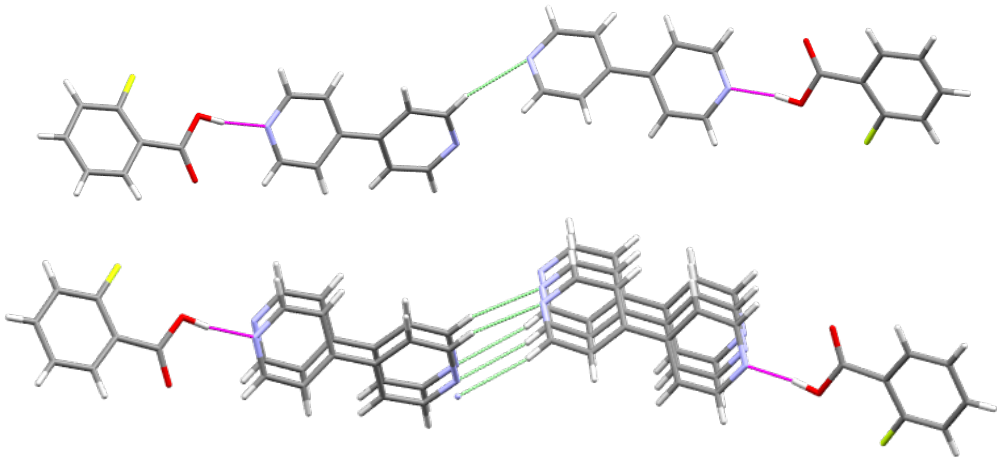


Figure 45: Hydrogen bonding motifs observed in 1:1 cocrystal 15.84 (hydrogen bonds in magenta and green).

4.6.2.4 Cocrystals of Urea [109]

In the case of urea [**109**], another commonly used coformer, there are 227 multi-component materials present in the CSD.[9] **109** is present in 63 cocrystals with carboxylic acid coformers, and in 42 cases (93% frequency) the acid-amide $R_2^2(8)$ dimer is observed. Due to the symmetrical diamide nature of urea, the amide-amide $R_2^2(8)$ homomeric dimer is also present in 30 structures (67% frequency), generally observed in conjunction with the acid amide dimer, in a symmetrical fashion as shown in Figure 46.

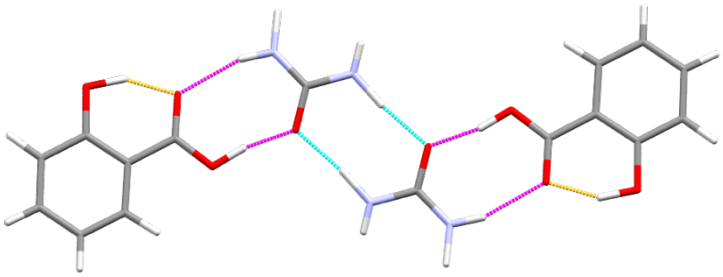


Figure 46: Combination of homo and heteromeric dimers in urea cocrystals, as observed in SLCADC01 [109.60].[37]

Of the 18 combinations available for cocrystal formation in Q1, **109** has successfully formed a cocrystal with 9 carboxylic acid coformers, of which 4 had not been previously described in the literature. Of the 5 known materials, the crystal structures of 2 were available in the CSD,[9] and the remaining 3 were described as cocrystals in the literature,[100,101] albeit without an accompanying crystal structure (Table 12).

Table 12: Cocrystallisation results in Q1 for 109 (novel solid forms identified in this study highlighted in green).

	Coformer	Space Group	CSD[9] Refcode
121	2-Nitrobenzoic Acid[a]	$P2_1/n$	N/A
122	3-Nitrobenzoic Acid[a]	$P2_1/c$	N/A
104	4-Nitrobenzoic Acid	-	No Structure[100]
60	2-Hydroxybenzoic Acid	$C2/c$	SLCADC/01/10[36–38]
82	3-Hydroxybenzoic Acid	-	No Structure[100]
83	4-Hydroxybenzoic Acid	-	No Structure[101]
84	2-Fluorobenzoic Acid	-	N/A
2	4-Aminobenzoic Acid	*Pnab*	NUHYEU[56]
79	2-Methylbenzoic Acid	-	N/A

a – Crystal structure determined using SCXRD;

Interestingly, in the cocrystal of urea and 4-aminobenzoic acid [**109.2**, NUHYEU],[56] the acid-amide $R_2^2(8)$ dimer does not form, as is observed in SLCADC01[37] [**109.60**] and many other cocrystals of this type. Instead, the urea molecules form a planar chain of homomeric $R_2^2(8)$ dimers along the centre of the structure, with capping of this chain occurring by hydrogen bonding from the free urea N-H heteromerically to the nearest carboxylic acid carbonyl. The acid molecules participate in a homomeric $R_2^2(8)$ dimer also, which run almost orthogonally to the urea chains (Figure 47). The urea carbonyl also accepts a

hydrogen bond from the *para* amino group, rendering it trifurcated in this case (Figure 48).

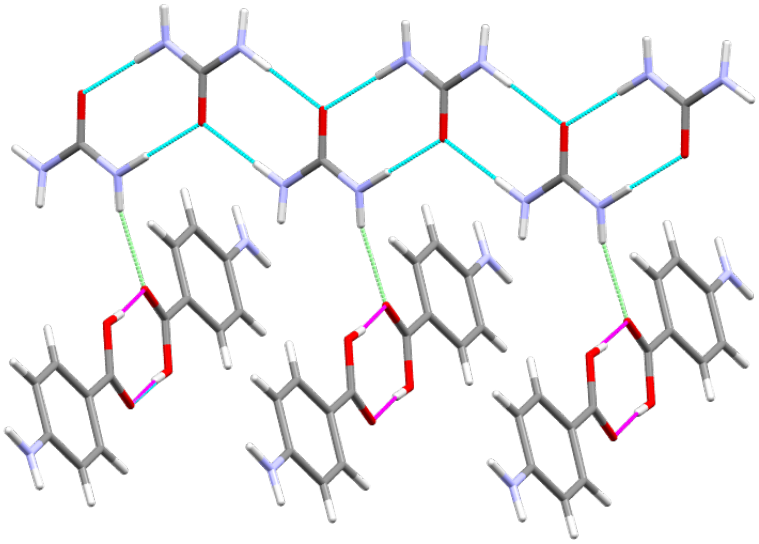


Figure 47: Primary hydrogen bonding interactions observed in 109.2 cocrystal (NUHYEU)[56] [homomeric dimers in magenta and cyan, heteromeric interactions in green].

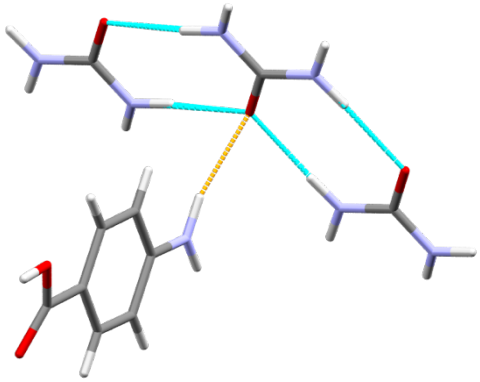


Figure 48: Trifurcation of the urea carbonyl in NUHYEU [109.2].[56]

The two novel cocrystals, **109.121** (2-nitrobenzoic acid) and **109.122** (3-nitrobenzoic acid), display similar motifs. Both form the heteromeric $R_2^2(8)$ acid-amide dimer, with the homomeric amide-amide dimer not observed in either case.

In the case of **109.121**, the material crystallizes in a 1:2 ratio, with preferential hydrogen bonding interactions occurring instead to the carboxylic acid, and the nitro groups. Each urea moiety is capped on all sides by carboxylic acid molecules, forming a 1-dimensional chain (Figure 49).

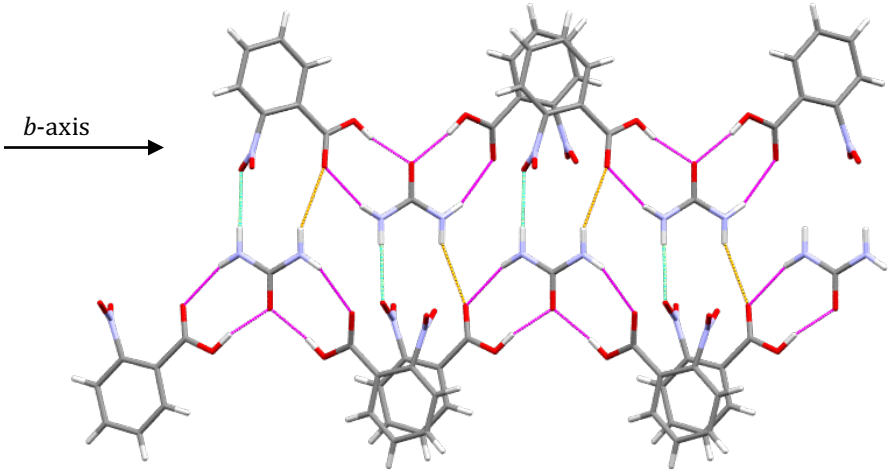


Figure 49: Hydrogen bonding interactions observed in 109.121.

The 1:1 cocrystal of **109.122** displays the dimer in an alternative fashion, capping an infinite C(4) chain of urea molecules that runs along the *c*-axis, creating a similar chain motif to that observed in **109.121** (Figure 50, magenta and orange). There is also interaction with the nitro groups in this case, capping the other side of the urea molecules (Figure 50, green). The remaining urea N-H_{3B} hydrogen bonds to a carboxylic acid carbonyl in a neighbouring layer (Figure 51, cyan), and there is also a nitro N=O···H-C bond observed here (Figure 51, green). The overall packing here is a complex set of hydrogen bonded layers, interlinked by the N-H···O=C hydrogen bonds between the **109** molecules.

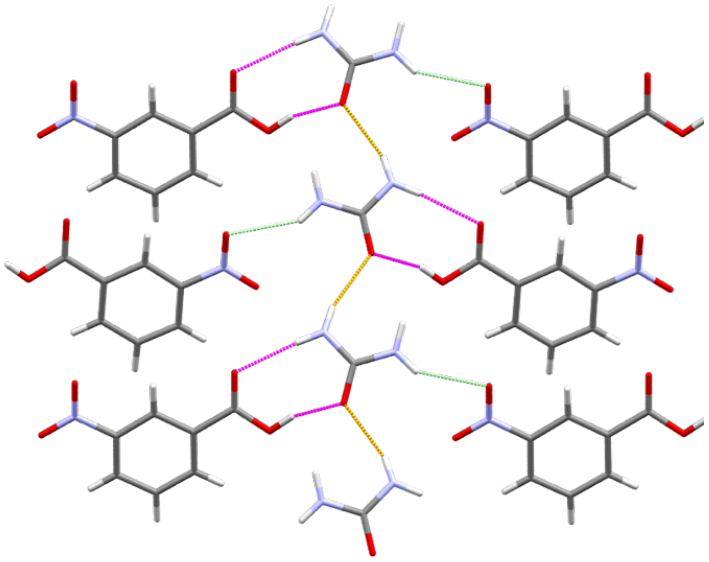


Figure 50: Hydrogen bonded dimers (magenta), C(4) chain (orange) and discrete nitro interactions (green) in 109.122.

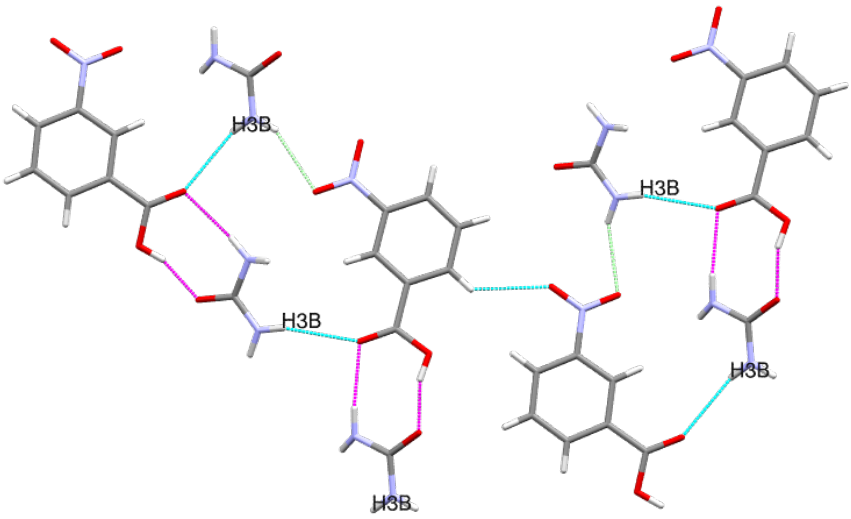


Figure 51: Additional hydrogen bonding interactions observed in 109.122 (hydrogen bonds in magenta, cyan and green).

4.6.2.5 Cocrystals of Benzamide [13]

Benzamide [**13**], the most simple aromatic amide used in this investigation, is contained in 24 cocrystal structures in the CSD, 13 of which contain a carboxylic acid coformer. The acid-amide $R_2^2(8)$ dimer is present in all cases here (100% frequency), and there is also high prevalence of the amide C(4) chain, present along with the dimer in 7 structures (Figure 52).

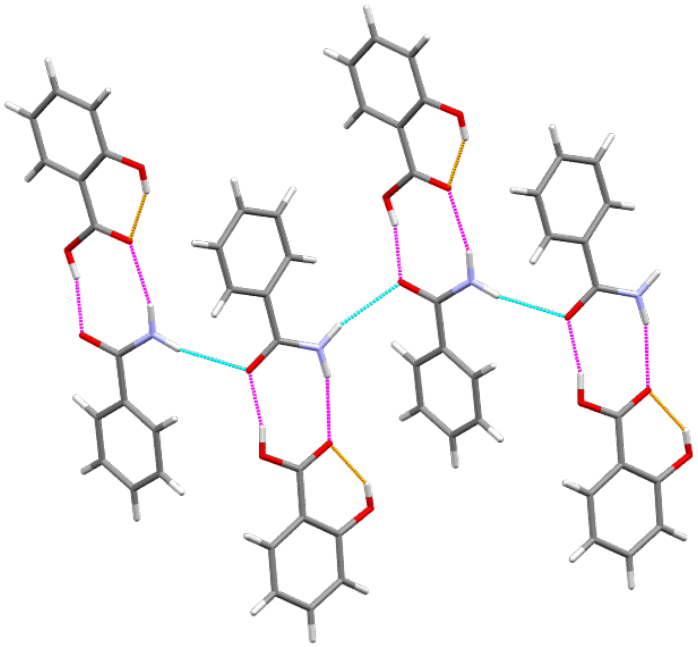


Figure 52: The dimer (magenta) and C(4) chain (cyan) observed in benzamide salicylic acid cocrystal 13.60 (URISAQ).[39]

In three structures, a heteromeric ladder-type motif is formed, similar to that observed for primary amides (Section 2.3.2). However, in this case the motif is not infinite. In two cases, OVEZUL[24] [**13.122**] and MEHCAF,[102] the ladder is formed via the interlinking of two acid-amide $R_2^2(8)$ dimers by N-H···O=C hydrogen bonds from the free amide N-H to the acid carbonyl, forming a binary $R_4^2(8)$ tetramer in the centre (Figure 53). Interestingly, in the case of QAFQEW,[34] a polymorphic form of **13.60** (polymorph of URISAQ/01),[24,39] the heteromeric ladder motif is not observed, but the dimer and C(4) chain combination can be seen; and separately, a trimeric ladder-type motif is displayed between the benzamide molecules (Figure 54, magenta, orange and cyan). In the third case, KEMCEL,[103] the homomeric amide-amide dimer is flanked by two heteromeric acid-amide dimers, creating a hexamer motif (Figure 55).

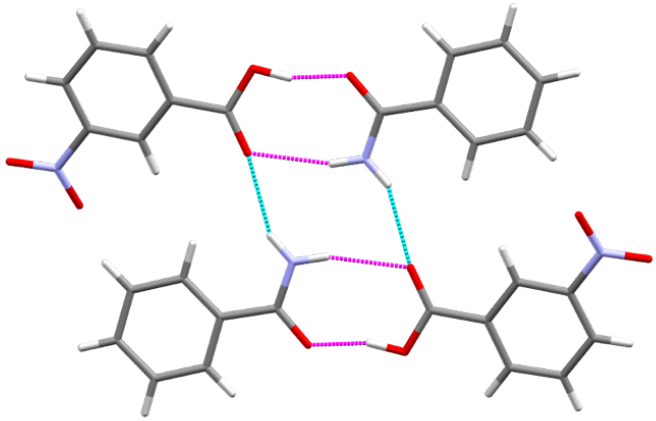


Figure 53: Ladder-type motif observed in cocrystal of 3-nitrobenzoic acid with benzamide [13.122, OVEZUL].[24]

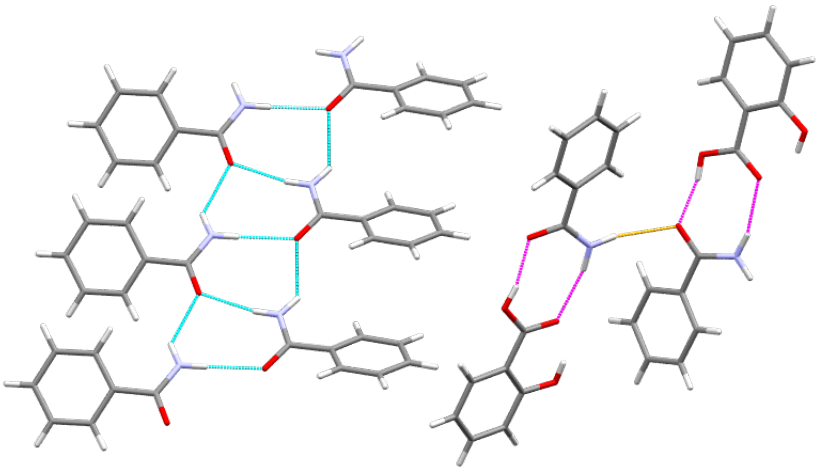


Figure 54: Motifs observed in polymorph of 13.60 (QAFQEW)[34] [heteromeric dimers in magenta, C(4) chain in orange and amide trimers in cyan].

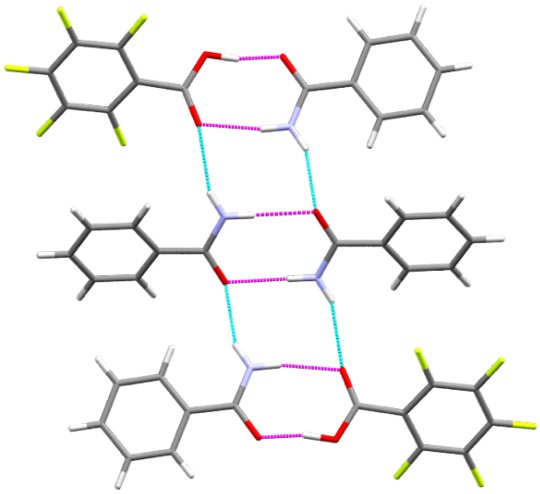


Figure 55: Extended ladder-type motif observed in KEMCEL, 2:1 cocrystal of benzamide with pentafluorobenzoic acid [13.132].[103]

Benzamide [**13**] was used successfully for cocrystal synthesis in 7 cases in Q1, of which 4 were novel solid forms. Two forms, **13.121** and **13.85**, were successfully characterised using SCXRD (Table 13).

Table 13: Cocrystallisation results in Q1 for 13 (novel solid forms identified in this study highlighted in green).

	Coformer	Space Group	CSD[9] Refcode
121	2-Nitrobenzoic Acid[a]	$P2_1/n$	N/A
122	3-Nitrobenzoic Acid[a]	$P2_1/c$	OVEZUL
104	4-Nitrobenzoic Acid[b]	P-1	YOPCAI
60	2-Hydroxybenzoic Acid	$P2_1/c$ $P2_12_12_1$	URISAQ/01 QAFQEW
84	2-Fluorobenzoic Acid	-	N/A
85	3-Fluorobenzoic Acid	P-1	N/A
86	4-Fluorobenzoic Acid	-	N/A

a – Crystal structure determined using SCXRD; b – Structural coordinates not determined;

The motifs discussed above are both represented in these novel crystal structures. **13.122** displays the ladder type motif observed in KEMCEL,[103] and OVEZUL, and **13.85** displays the combination of $R_2^2(8)$ dimers and C(4) chain shown in URISAQ/01[24,39] and QAFQEW [**13.60**].[34]

13.121 crystallizes in a 2:1 ratio, with 2 unique molecules of 2-nitrobenzoic acid in the structure. The $R_4^2(8)$ tetramer motifs created by the interaction of the acide-amide dimers (Figure 56) are linked together *via* C-H···O=N hydrogen bonds to the nitro groups (Figure 57).

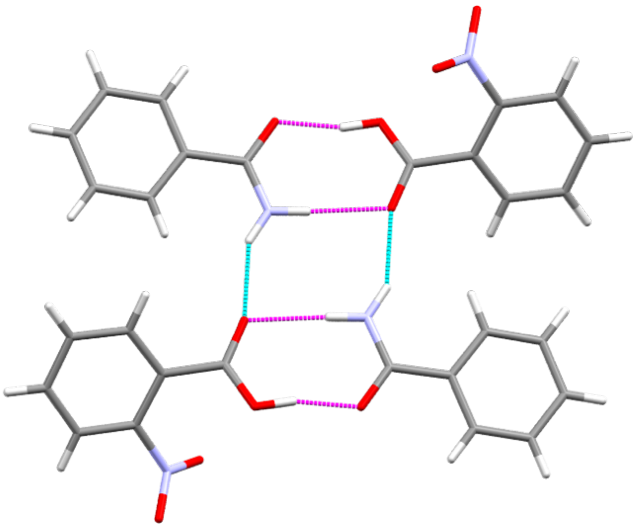


Figure 56: Tetramer motif observed in 13.121

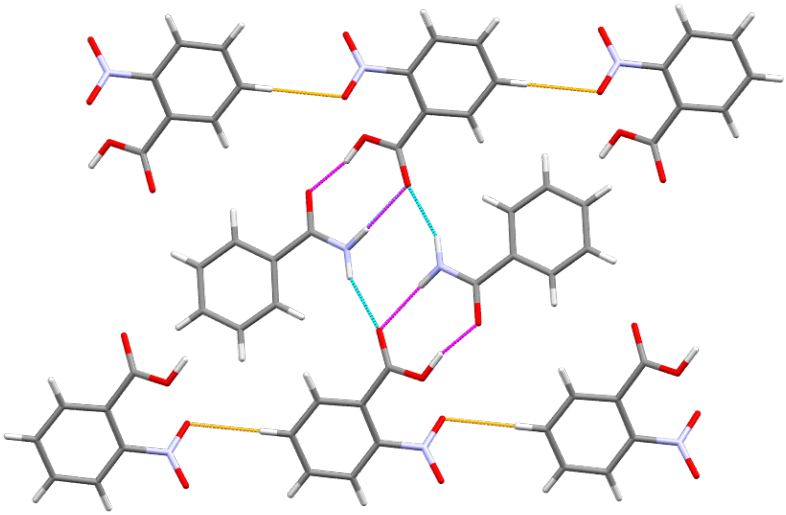


Figure 57:Hydrogen bonding interactions observed in 13.121 [heteromeric dimer in magenta, tetramer in cyan/magenta, C-H···O=N interactions in orange]

The hydrogen bonds to the nitro group are structurally significant here, aligning the heteromeric $R_2^2(8)$ dimers along the *a*-axis. The benzoic acid materials also interact homomerically here, creating the 2:1 ratio observed, and these homomeric dimers are linked to the dominant heteromeric motifs *via* a system of weak C-H···O hydrogen bonds (Figure 58).

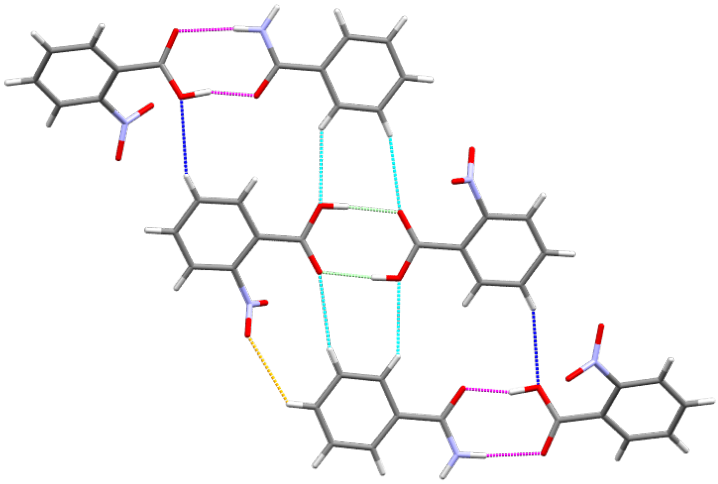


Figure 58: Weak C-H···O hydrogen bonds (cyan, orange and blue) linking homomeric (green) and heteromeric (magenta) $R_2^2(8)$ dimer motifs in 13.121.

The complexity of the interactions observed here contrasts significantly with that observed in **13.85**, which displays the simple combination of $R_2^2(8)$ dimer and C(4) chain as discussed above (Figure 52). These chains are interlinked by similar weak (2.7 Å) C-H···O hydrogen bonds, similar to those observed in **13.121** (Figure 59).

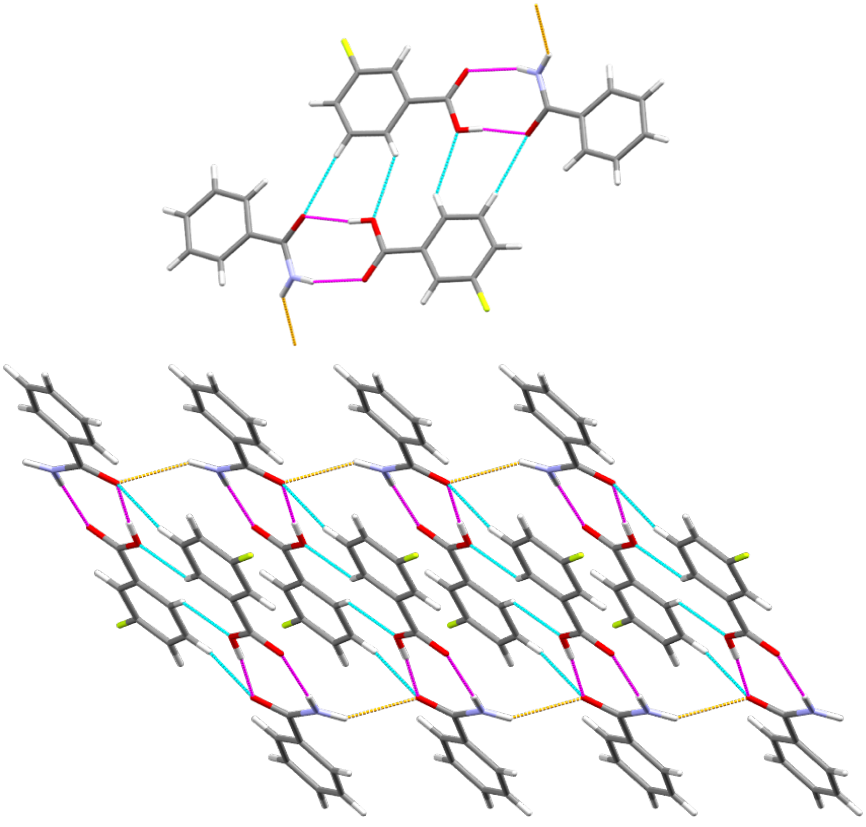


Figure 59: Hydrogen bonding motifs observed in 13.85 [dimer in magenta, C(4) chain in orange, and C-H···O bonds in cyan].

4.7 Conclusions

Generally, the cocrystal structures determined in this study display a high level of consistency with the anticipated hydrogen bonded motifs. Despite the prevalence of the primary hydrogen bonding targets (Section 1.1.4), there is significant structural variation observed between these materials, driven by secondary interactions to the substituted aromatic rings. The results of this study highlight the synthetic utility of the four primary hydrogen bonding targets in cocrystal synthesis, while concurrently exposing the level of structural variation that can be observed within systems of similar molecules.

The level of structural variation observed here highlights the utility of an efficient cocrystal ranking tool such as that developed in this study. A combinatorial approach using both knowledge based cocrystal design and computational prediction provides significant advantages over traditional high throughput screening by increasing the success rate and minimizing number of unsuccessful outcomes thereby providing a tangible saving in both time and resources. The simplicity of the model developed in this study allows for its use in many scenarios, with the possibility to predict outcomes for a huge range of target materials with the given set of acid and amide coformers. The approach used here could be tailored or expanded for use in an industrial setting, for example, the development of a data matrix of suitable GRAS coformers for cocrystallization prediction with APIs.

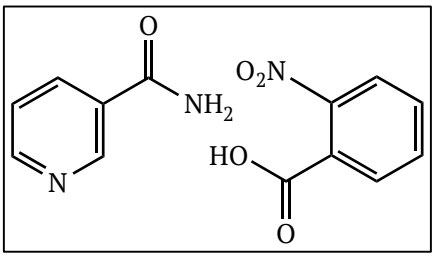
Chapter 4

Experimental

4.8 Experimental

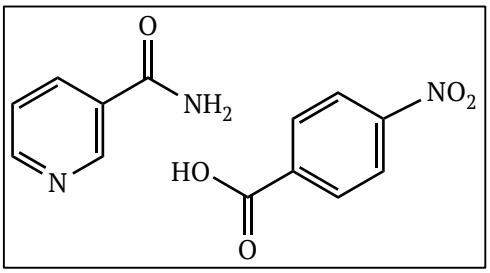
4.8.1 Cocrystals of Nicotinamide [1]

Nicotinamide 2-nitrobenzoic acid cocrystal [1.121]



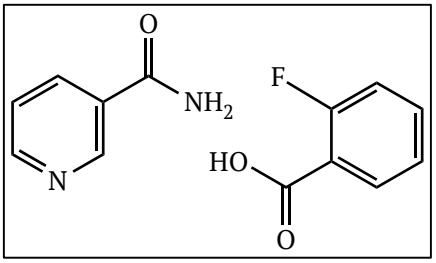
2-Nitrobenzoic acid [**121**] (0.0418 g, 0.25 mmol) and nicotinamide [**1**] (0.0305 g, 0.25 mmol) were used. Colourless block crystals of **1.121** were obtained from acetonitrile; DSC (endotherm): 87 °C, 98 °C, 109-114 °C; mp 90-93 °C; ν_{max} (ATR)/cm^{-1}; 3378, 3166 (N-H stretches), 1689 (C=O), 1583 (Aromatic C=C bend), 1356 (NO_2 stretch), 1050, 813, 783; Crystal data for **1.121**: $C_{20}H_{16}O_9N_4$, *Mr* = 456.37, monoclinic, *C*2/c, *a* = 27.715(3) Å, *b* = 7.0371(7) Å , *c* = 21.947(2) Å, β = 105.132(4)°, V = 4132.0(7) Å^3, *Z* = 8, D_c = 1.467 g cm^{-3}, F_{000} = 1888, Mo Kα radiation, λ = 0.71073 Å, *T* = 300(2) K, $2\theta_{max}$ = 25.01°, μ = 0.118 mm^{-1}, 22442 reflections collected, 3633 unique (R_{int} = 0.0431), final GooF = 1.019, R_1 = 0.0401 [2875 obs. data: $I>2\sigma(I)$], wR_2 = 0.1092 (all data).

Nicotinamide 4-nitrobenzoic acid cocrystal [1.104]



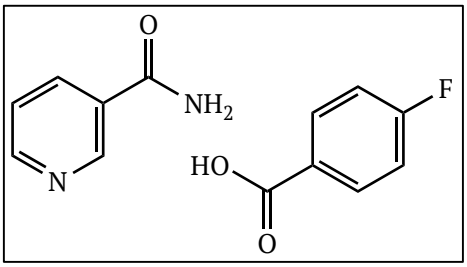
4-Nitrobenzoic acid [**104**] (0.0418 g, 0.25 mmol) and nicotinamide [**1**] (0.0305 g, 0.25 mmol) were used. Colourless block crystals of **1.104** were obtained from acetonitrile; DSC (endotherm): 171.5-172.5 °C; ν_{max} (ATR)/cm^{-1}; 3396 (O-H stretch), 3376, 3167 (N-H stretches), 1682 (C=O), 1601, 1519 (Aromatic C=C bend), 1313 (NO_2 stretch), 1010, 794, 720; Crystal data for **1.104**: $C_{13}H_{11}O_5N_3$, *Mr* = 289.25, triclinic, *P*-1, *a* = 7.1167(5) Å, *b* = 7.5590(5) Å , *c* = 12.8081(9) Å, α = 85.164(2)°, β = 75.933(2)°, γ = 85.895(2)°, *V* = 665.04(8) Å^3, *Z* = 2, D_c = 1.444 g cm^{-3}, F_{000} = 300, Mo Kα radiation, λ = 0.71073 Å, *T* = 296.(2) K, $2\theta_{max}$ = 26.37°, μ = 0.114 mm^{-1}, 19325 reflections collected, 2718 unique (R_{int} = 0.0344), final GooF = 1.023, R_1 = 0.0397 [2064 obs. data: $I>2\sigma(I)$], wR_2 = 0.1159 (all data).

Nicotinamide 2-fluorobenzoic acid cocrystal [1.84]



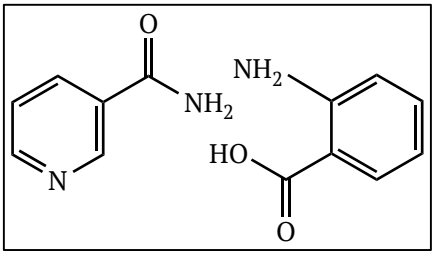
2-Fluorobenzoic acid [**84**] (0.0352 g, 0.25 mmol) and nicotinamide [**1**] (0.0305 g, 0.25 mmol) were used. Colourless block crystals of **1.84** were obtained from acetonitrile; DSC (endotherm): 93-96 °C; ν_{max} (ATR)/cm^{-1}; 3379, 3203 (N-H stretches), 1638 (C=O), 1610 (C=O), 1395, 1295 (C-F), 870, 755; Single crystal data was collected for this material, confirming the formation of a 1:1 cocrystal. However, the structure obtained presented a high degree of disorder around the fluoro-phenyl ring and the R-factor was too high for the structure to be included in this thesis [$P2_1/c$, a = 19.480(6), b = 4.9935(16), c = 12.872(4) Å, β = 91.888(7)°].

Nicotinamide 4-fluorobenzoic acid cocrystal [1.86]



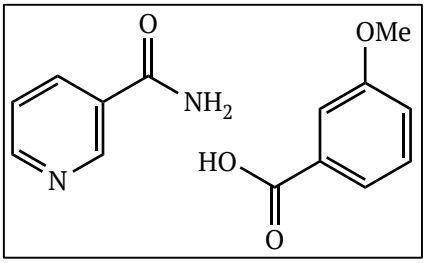
4-Fluorobenzoic acid [**86**] (0.0352 g, 0.25 mmol) and nicotinamide [**1**] (0.0305 g, 0.25 mmol) were used. Colourless block crystals of **1.86** were obtained from acetonitrile; DSC (endotherm): 141-142 °C; ν_{max} (ATR)/cm^{-1}; 3406 (O-H stretch), 3166 (N-H stretch), 1671 (C=O), 1600 (C=O), 1398, 1244, 1224, 865, 768; Crystal data for **1.86**: $C_{13}H_{11}O_3N_2F$, Mr = 262.24, monoclinic, $P2_1/c$, a = 13.629(16) Å, b =7.151(8) Å , c = 13.651(15) Å, β = 115.649(17)°, V = 1199.0(2) Å^3, Z = 4, D_c = 1.452 g cm^{-3}, F_{000} = 544, Mo Kα radiation, λ = 0.71073 Å, T = 296.(2) K, $2\theta_{max}$ = 26.52°, μ = 0.115 mm^{-1}, 6819 reflections collected, 2463 unique (R_{int} = 0.0286), final GooF = 1.043, R_1 = 0.0390 [1764 obs. Data: $I>2\sigma(I)$], wR_2 = 0.1090 (all data).

Nicotinamide 2-aminobenzoic acid cocrystal [1.77]



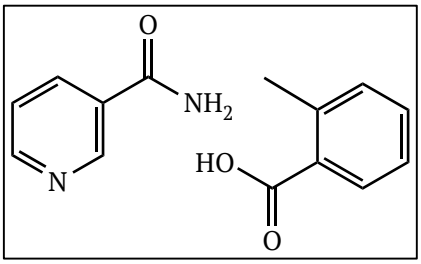
2-Aminobenzoic acid [**77**] (0.0343 g, 0.25 mmol) and nicotinamide [**1**] (0.0305 g, 0.25 mmol) were used. Colourless needles of **1.77** were obtained from ethanol; DSC (endotherm): 110-114 °C; ν_{max} (ATR)/cm^{-1}; 3436 (w) (O-H stretch), 3392, 3330, 3218 (w) (N-H stretches), 1657 (C=O), 1296, 1196 (C-O); Crystal data for **1.77**: $C_{13}H_{11}N_3O_3$, *Mr* = 259.26, monoclinic, $P2_1$, a = 10.479(2) Å, b =4.9873(9) Å , c = 12.644(3) Å, β = 109.361(5)°, *V* = 623.4(2) Å^3, *Z* = 2, D_c = 1.381 g cm^{-3}, F_{000} = 272, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 25.68°, μ = 0.101 mm^{-1}, 9087 reflections collected, 2364 unique (R_{int} = 0.0278), final GooF = 1.040, R_1 = 0.0305 [2141 obs. Data: $I>2\sigma(I)$], wR_2 = 0.0702 (all data).

Nicotinamide 3-methoxybenzoic acid cocrystal [1.124]



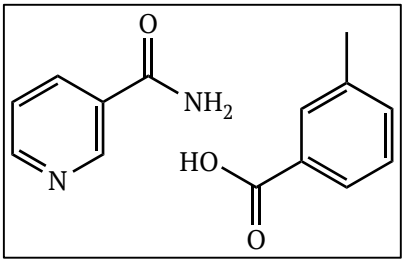
3-Methoxybenzoic Acid [**124**] (0.0382 g, 0.25 mmol) and nicotinamide [**1**] (0.0305 g, 0.25 mmol) were used. DSC (endotherm): 97-99 °C; ν_{max} (ATR)/cm^{-1}; 3458 (w) (O-H stretch), 3373, 3160 (w) (N-H stretches), 1674 (C=O), 1580 (Aromatic C=C bend), 1283, 1200 (C-O), 1039, 754.

Nicotinamide 2-methylbenzoic acid cocrystal [1.79]



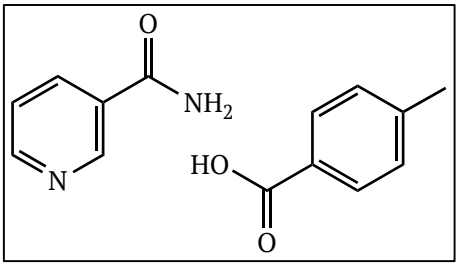
2-Methylbenzoic Acid [**79**] (0.0341 g, 0.25 mmol) and nicotinamide [**1**] (0.0305 g, 0.25 mmol) were used. DSC (endotherm): 78-80 °C; ν_{max} (ATR)/cm^{-1}; 3394, 3147 (w) (N-H stretches), 1712 (C=O), 1668 (C=O), 1600, 1393, 1297, 1267, 1044, 732.

Nicotinamide 3-methylbenzoic acid cocrystal [1.80]



3-Methylbenzoic Acid [**80**] (0.0680 g, 0.5 mmol) and nicotinamide [**1**] (0.0305 g, 0.25 mmol) were used. Colourless block crystals of **1.80** were obtained from ethanol; DSC (endotherm): 97 °C, 102 °C, and 107-109 °C; mp 122-124 °C; ν_{max} (ATR)/cm^{-1}; 3361, 3183 (w) (N-H stretches), 1800 (C=O), 1373, 1213, 746; Crystal data for **1.80**: $C_{22}H_{22}N_2O_5$, *Mr* = 394.41, monoclinic, $P2_1/n$, *a* = 10.878(3) Å, *b* = 12.704(4) Å , *c* = 15.328(5) Å, β = 107.034(9)°, *V* = 2025.3(11) Å^3, *Z* = 4, D_c = 1.294 g cm^{-3}, F_{000} = 832, Mo Kα radiation, λ = 0.71073 Å, *T* = 300(2) K, $2\theta_{max}$ = 25.06°, μ = 0.093 mm^{-1}, 37173 reflections collected, 3550 unique (R_{int} = 0.1220), final GooF = 1.036, R_1 = 0.0672 [1651 obs. ata: $I>2\sigma(I)$], wR_2 = 0.1559 (all data).

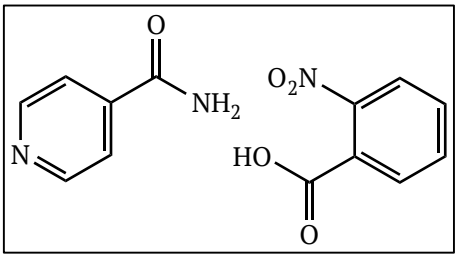
Nicotinamide 4-methylbenzoic acid cocrystal [1.81]



4-Methylbenzoic Acid [**81**] (0.0340 g, 0.25 mmol) and nicotinamide [**1**] (0.0305 g, 0.25 mmol) were used. DSC (endotherm): 83-85 °C; ν_{max} (ATR)/cm^{-1}; 3373, 3181 (w) (N-H stretches), 1645 (C=O), 1308, 1287, 754. Single crystals of this material suitable for SCXRD could not be prepared.

4.8.2 Cocrystals of Isonicotinamide [61]

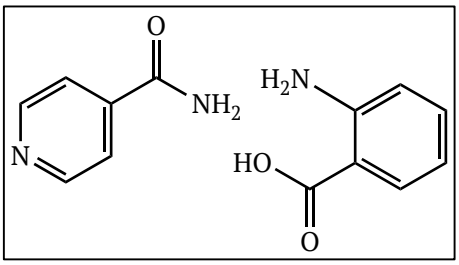
Isonicotinamide 2-nitrobenzoic acid cocrystal [61.121]



2-Nitrobenzoic acid [**121**] (0.0418 g, 0.25 mmol) and isonicotinamide [**61**] (0.0305 g, 0.25 mmol) were used. Colourless brick crystals of **61.121** were obtained from acetonitrile; DSC (endotherm): 126-129 °C; ν_{max} (ATR)/cm^{-1}; 1709 (C=O), 1577 (Aromatic C=C bend), 1355 (NO_2 stretch),

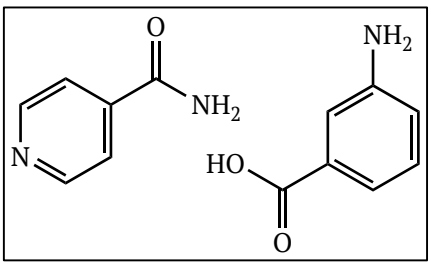
739, 705; Crystal data for **61.121**: $C_{13}H_{11}N_3O_5$, *Mr* = 289.25, monoclinic, $P2_1/c$, *a* = 8.873(2) Å, *b* = 34.245(8) Å , *c* = 9.175(2) Å, β = 105.942(8)°, *V* = 2680.7(11) Å^3, *Z* = 8, D_c = 1.433 g cm^{-3}, F_{000} = 1200, Mo Kα radiation, λ = 0.71073 Å, *T* = 300(2) K, $2\theta_{max}$ = 25.85°, μ = 0.113 mm^{-1}, 29445 reflections collected, 4734 Unique (R_{int} = 0.0540), final GooF = 1.095, R_1 = 0.0482 [3588 obs. data: $I>2\sigma(I)$], wR_2 = 0.1668 (all data).

Isonicotinamide 2-aminobenzoic acid cocrystal [61.77]



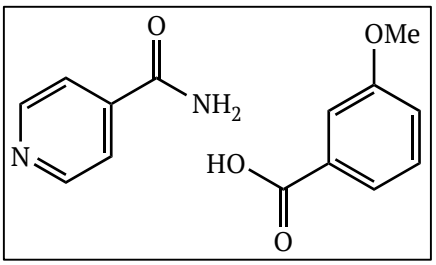
2-Aminobenzoic acid [**77**] (0.0343 g, 0.25 mmol) and isonicotinamide [**61**] (0.0305 g, 0.25 mmol) were used. Orange plate crystals of **61.77** were obtained from acetonitrile; DSC (endotherm): 113-115.5 °C; ν_{max} (ATR)/cm^{-1}; 3432 (w) (O-H stretch), 3330, 3156 (w) (N-H stretches), 1680 (C=O), 1297, 1241, 1228 (C-O), 753; Crystal data for **61.77**: $C_{13}H_{13}N_3O_3$, *Mr* = 259.26, monoclinic, $P2_1/c$, *a* = 12.516(5) Å, *b* = 10.899(4) Å , *c* = 9.306(3) Å, β = 95.296(12)°, *V* = 1264.0(8) Å^3, *Z* = 4, D_c = 1.362 g cm^{-3}, F_{000} = 544, Mo Kα radiation, λ = 0.71073 Å, *T* = 300(2) K, $2\theta_{max}$ = 25.03°, μ = 0.099 mm^{-1}, 19310 reflections collected, 2225 Unique (R_{int} = 0.1923), final GooF = 1.015, R_1 = 0.0853 [1007 obs. data: $I>2\sigma(I)$], wR_2 = 0.2536 (all data).

Isonicotinamide 3-aminobenzoic acid cocrystal [61.78]



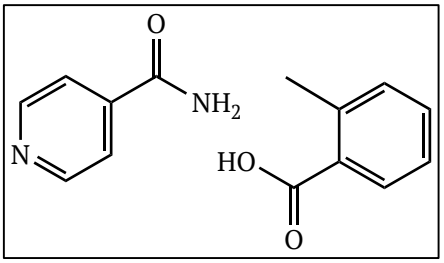
3-Aminobenzoic acid [**78**] (0.0343 g, 0.25 mmol) and isonicotinamide [**61**] (0.0305 g, 0.25 mmol) were used. DSC (endotherm): 107-109 °C; ν_{max} (ATR)/cm^{-1}; 3429 (w) (O-H stretch), 3364, 3322, 3175 (w) (N-H stretches), 1697 (C=O), 1555 (Aromatic C=C bend), 1463 (Aromatic C=C bend), 1410, 1254, 1209 (C-O), 1178 (C-O), 1016, 751.

Isonicotinamide 3-methoxybenzoic acid cocrystal [61.124]



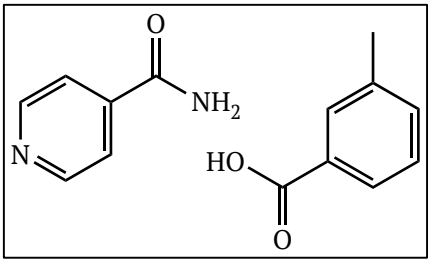
3-Methoxybenzoic acid [**124**] (0.0381 g, 0.25 mmol) and isonicotinamide [**61**] (0.0305 g, 0.25 mmol) were used. Colourless block crystals of **61.124** were obtained from acetonitrile; DSC (endotherm): 130-133 °C; ν_{max} (ATR)/cm^{-1}; 3354, 3162 (N-H stretches), 1687 (C=O), 1582 (Aromatic C=C bend), 1243, 1229 (C-O), 1121 (C-O), 1042, 753; Crystal data for **61.124**: $C_{44}H_{44}O_{14}N_4$, *Mr* = 852.84, triclinic, *P*-1, *a* = 10.443(8) Å, *b* = 12.603(11) Å , *c* = 17.396(16) Å, α = 110.90(2)°, β = 98.50(2)°, γ = 90.185(19)°, *V* = 2112(3) Å^3, *Z* = 6, D_c = 1.341 g cm^{-3}, F_{000} = 896, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 26.92°, μ = 0.101 mm^{-1}, 36421 reflections collected, 8786 Unique (R_{int} = 0.0494), final GooF = 1.018, R_1 = 0.0659 [4679 obs. data: $I>2\sigma(I)$], wR_2 = 0.2089 (all data).

Isonicotinamide 2-methylbenzoic acid cocrystal [61.79]



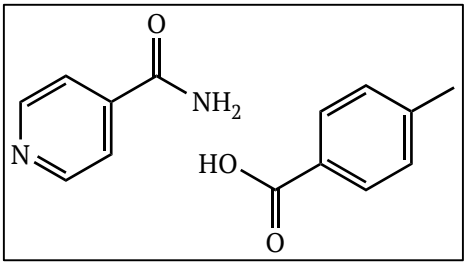
2-Methylbenzoic acid [**79**] (0.0340 g, 0.25 mmol) and isonicotinamide [**61**] (0.0305 g, 0.25 mmol) were used. DSC (endotherm); 108-110 °C, 116-117 °C, 119-121 °C, 134-135.5 °C; mp 108-110 °C; ν_{max} (ATR)/cm^{-1}; 3379, 3164 (w) (N-H stretches), 1671 (C=O), 1408 (Aromatic C=C bend), 1229 (C-O), 737.

Isonicotinamide 3-methylbenzoic acid cocrystal [61.80]



3-Methylbenzoic acid [**80**] (0.0340 g, 0.25 mmol) and isonicotinamide [**61**] (0.0305 g, 0.25 mmol) were used. DSC (endotherm) 135-136 °C; ν_{max} (ATR)/cm^{-1}; 3349, 3165 (w) (N-H stretches), 1743 (C=O), 1693 (C=O), 1410 (Aromatic C=C bend), 1231, 746.

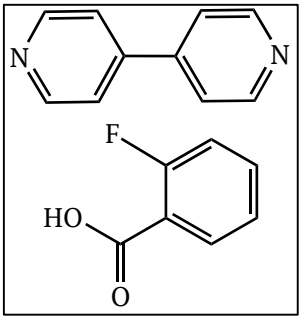
Isonicotinamide 4-methylbenzoic acid cocrystal [61.81]



4-Methylbenzoic acid [**81**] (0.0680 g, 0.50 mmol) and isonicotinamide [**61**] (0.0305 g, 0.25 mmol) were used. Colourless block crystals of **61.81** were obtained from acetonitrile. DSC (endotherm): 181-183 °C , 185-187 °C; mp 182-184 °C; ν_{max} (ATR)/cm^{-1}; 3366, 3177 (w) (N-H stretches), 1668 (C=O), 1611 (C=O), 1282, 753; Single Crystal data was collected for this material, confirming the formation of a 1:1 cocrystal. However, the R-factor was too high for the structure to be included in this thesis [*P-1, a* = 10.078(2), *b* = 12.583(3), *c* = 16.235(4) Å, α = 76.239(6)°, *β* = 81.037(5)°, γ = 89.860(5)°].

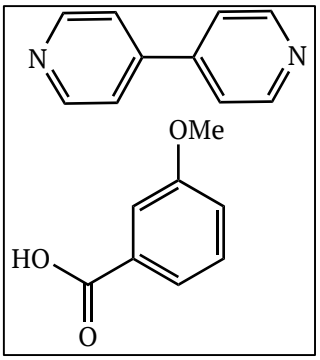
4.8.3 Cocrystals of 4,4'-Bipyridyl [15]

4,4'-Bipyridyl 2-fluorobenzoic acid cocrystal [15.84]



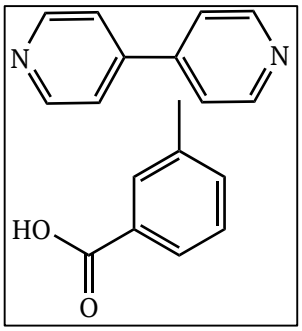
2-Fluorobenzoic acid [**84**] (0.0350 g, 0.25 mmol) and 4,4'-bipyridyl [**15**] (0.0392 g, 0.25 mmol) were used. Colourless block crystals of **15.84** were obtained from ethanol. DSC (endotherm): 88-90 °C (minor), 101-103 °C (major); mp 101-103 °C; ν_{max} (ATR)/cm^{-1}; 1692 (C=O), 1591 (Aromatic C=C bend), 1534 (Aromatic C=C bend), 1406, 1302, 1246 (C-F), 1148 (C-O), 873, 843; Crystal data for **15.84**: $C_{17}H_{13}O_2N_2F$, *Mr* = 296.29, monoclinic, $P2_1/n$, *a* = 11.012(2) Å, *b* = 4.0528(8) Å , *c* = 32.335(6) Å, *β* = 94.648(4)°, *V* = 1438.3(5) Å^3, *Z* = 4, D_c = 1.368 g cm^{-3}, F_{000} = 616, Mo Kα radiation, λ = 0.71073 Å, *T* = 296.(2) K, $2\theta_{max}$ = 26.46°, μ = 0.100 mm^{-1}, 21090 reflections collected, 2944 Unique (R_{int} = 0.0370), final GooF = 1.044, R_1 = 0.0589 [2002 obs. Data: $I>2\sigma(I)$], wR_2 = 0.2075 (all data).

4,4'-Bipyridyl 3-methoxybenzoic acid cocrystal [15.124]



3-Methoxybenzoic acid [**124**] (0.0382 g, 0.25 mmol) and 4,4'-bipyridyl [**15**] (0.0392 g, 0.25 mmol) were used. DSC (endotherm) 122-124 °C; ν_{max} (ATR)/cm^{-1}; 1695 (C=O), 1609, 1407 (Aromatic C=C bend), 1305, 1294, 1275, 1214 (C-O), 1047, 989, 809, 749.

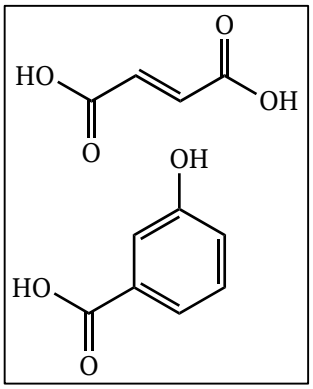
4,4'-Bipyridyl 3-methylbenzoic acid cocrystal [15.80]



3-Methylbenzoic acid [**80**] (0.0340 g, 0.25 mmol) and 4,4'-bipyridyl [**15**] (0.0392 g, 0.25 mmol) were used. DSC (endotherm) 92-94 °C; ν_{max} (ATR)/cm^{-1}; 1699 (C=O), 1585 (Aromatic C=C bend), 1471 (Aromatic C=C bend), 924, 877, 756.

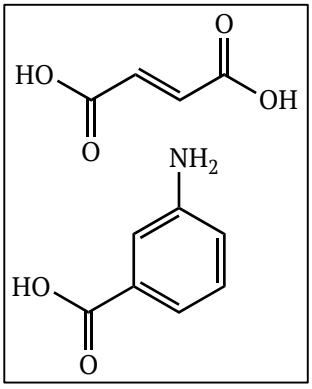
4.8.4 Cocrystals of Fumaric Acid [108]

Fumaric acid 3-hydroxybenzoic acid [108.82]



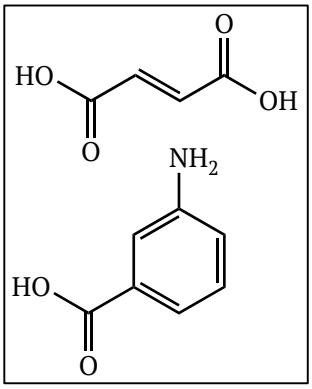
3-Hydroxybenzoic acid [**82**] (0.0345 g, 0.25 mmol) and fumaric acid [**108**] (0.0290 g, 0.25 mmol) were used. DSC (endotherm) 194-196 °C; ν_{max} (ATR)/cm^{-1}; 3334 (w) (O-H), 1668 (C=O), 1599 (Aromatic C=C bend), 1425 (Aromatic C=C bend), 1307, 1274, 1256, 1230 (C-O), 919, 900, 885, 756.

Fumaric acid 3-aminobenzoic acid [108.78]



3-Aminobenzoic acid [**78**] (0.0343 g, 0.25 mmol) and fumaric acid [**108**] (0.0290 g, 0.25 mmol) were used. DSC (endotherm) 172-178 °C; mp 177-178 °C; ν_{max} (ATR)/cm^{-1}; 1682 (C=O), 1642 (C=O), 1589 (Aromatic C=C bend), 1459 (Aromatic C=C bend), 1421 (Aromatic C=C bend), 1275, 1223, 1213 (C-O), 956, 767.

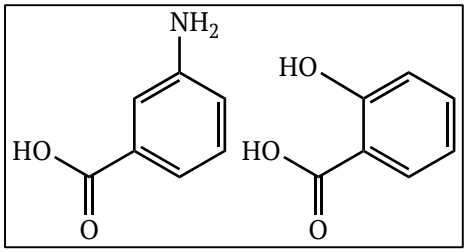
Fumaric acid 2-methylbenzoic acid [108.79]



2-Methylbenzoic acid [**79**] (0.0340 g, 0.25 mmol) and fumaric acid [**108**] (0.0290 g, 0.25 mmol) were used. mp 145-149 °C; ν_{max} (ATR)/cm^{-1}; 1682 (C=O), 1642 (C=O), 1589 (Aromatic C=C bend), 1459 (Aromatic C=C bend), 1421 (Aromatic C=C bend), 1275, 1223 (C-O), 1213 (C-O), 956, 767.

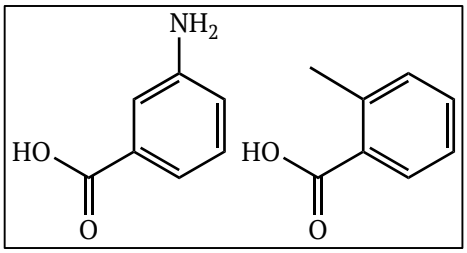
4.8.5 Cocrystals of Salicylic Acid [60]

2-Hydroxybenzoic acid 3-aminobenzoic acid cocrystal [60.78]



3-Aminobenzoic acid [**78**] (0.0343 g, 0.25 mmol) and 2-hydroxybenzoic acid [**60**] (0.0345 g, 0.25 mmol) were used. DSC (endotherm) 128-130 °C; ν_{max} (ATR)/cm^{-1}; 3236 (w) (N-H stretch), 1656 (C=O), 1614 (C=0), 1561 (Aromatic C=C bend), 1482 (Aromatic C=C bend), 1443 (Aromatic C=C bend), 1381, 1248, 1209 (C-O), 1155, 757.

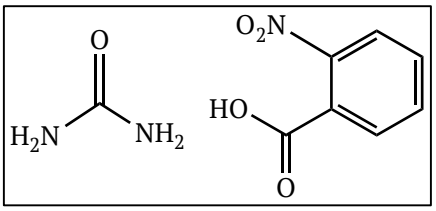
2-Hydroxybenzoic acid 2-methylbenzoic acid cocrystal [60.79]



2-Methylbenzoic acid [**79**] (0.0340 g, 0.25 mmol) and 2-hydroxybenzoic acid [**60**] (0.0345 g, 0.25 mmol) were used. DSC (endotherm) 92.5 °C (minor), 94-96 °C (major); mp 93-97°C; ν_{max} (ATR)/cm^{-1}; 3236 (w) (N-H stretch), 1657 (C=O), 1611 (C=0), 1483 (Aromatic C=C bend), 1382, 1295, 1270, 1209 (C-O), 758, 736.

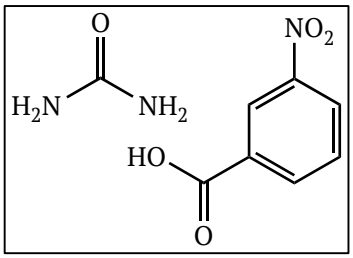
4.8.6 Cocrystals of Urea [109]

Urea 2-nitrobenzoic acid cocrystal [109.121]



2-Nitrobenzoic acid [**121**] (0.0835 g, 0.50 mmol) and urea [**109**] (0.0152 g, 0.25 mmol) were used. Colourless block crystals of **109.121** were obtained from acetonitrile. DSC (endotherm): 63-67 °C; ν_{max} (ATR)/cm^{-1}; 3467, 3433 (w) (N-H stretches), 1687 (C=O), 1619 (C=O), 1488 (Aromatic C=C bend), 1425 (Aromatic C=C bend), 1328 (N=O), 1279, 1260, 772. Crystal data for **109.121**: $C_{15}H_{14}O_9N_4$, *Mr* = 394.30, monoclinic, $P2_1/n$, a = 11.8242(18) Å, b = 10.0350(15) Å, c = 15.060(2) Å, β = 104.953(2)°, V = 1726.4(4) Å^3, Z = 4, D_c = 1.517 g cm^{-3}, F_{000} = 816, Mo Kα radiation, λ = 0.71073 Å, T = 296.(2) K, $2\theta_{max}$ = 25.14°, μ = 0.128 mm^{-1}, 22669 reflections collected, 3077 Unique (R_{int} = 0.0288), final GooF = 1.038, R_1 = 0.0366 [2530 obs. Data: $I>2\sigma(I)$], wR_2 = 0.1012 (all data).

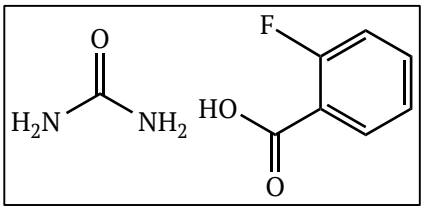
Urea 3-nitrobenzoic acid cocrystal [109.122]



3-Nitrobenzoic acid [**122**] (0.0418 g, 0.25 mmol) and urea [**109**] (0.0152 g, 0.25 mmol) were used. Colourless brick crystals of **109.122** were obtained from ethanol. DSC (endotherm) 149-152 °C; ν_{max}

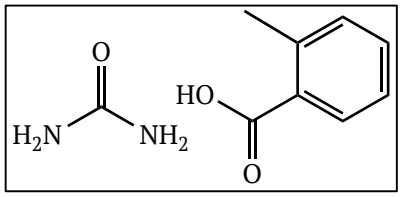
(ATR)/cm^{-1}; 3349, 3165 (w) (N-H stretches), 1743 (C=O), 1693 (C=O), 1410 (Aromatic C=C bend), 1350 (N=O), 1231, 746; Crystal data for **109.122**: $C_{15}H_{14}O_9N_4$, *Mr* = 394.30, monoclinic, $P2_1/c$, *a* = 8.084(4) Å, *b* = 12.756(6) Å, *c* = 9.490(4) Å, β = 93.543(12)°, *V* = 966.7(8) $Å^3$, *Z* = 4, D_c = 1.545 g cm^{-3}, F_{000} = 472, Mo Kα radiation, λ = 0.71073 Å, *T* = 296(2) K, $2\theta_{max}$ = 26.69°, μ = 0.131 mm^{-1}, 7065 reflections collected, 1139 unique (R_{int} = 0.0580), final GooF = 1.126, R_1 = 0.0642 [1904 obs. data: $I>2\sigma(I)$], wR_2 = 0.2149 (all data).

Urea 2-fluorobenzoic acid cocrystal [109.84]



2-Fluorobenzoic acid [**84**] (0.0835 g, 0.50 mmol) and urea [**109**] (0.0152 g, 0.25 mmol) were used. DSC (endotherm) 83-85 °C; ν_{max} (ATR)/cm^{-1}; 3446, 3319 (w) (N-H stretches), 1700 (C=O), 1646 (C=O), 1612, 1264 (C-F), 841.

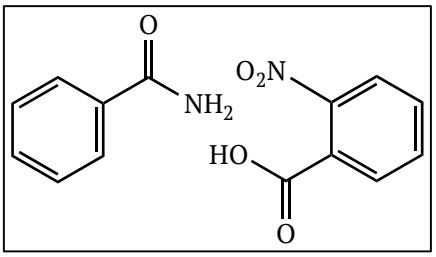
Urea 2-methylbenzoic acid cocrystal [109.79]



2-Methylbenzoic acid [**79**] (0.0340 g, 0.25 mmol) and urea [**109**] (0.0152 g, 0.25 mmol) were used. DSC (endotherm) 85 °C (minor), 102-104 °C (major); mp 103-105 °C; ν_{max} (ATR)/cm^{-1}; 3432, 3335 (w) (N-H stretches), 1678 (C=O), 1621 (C=O), 1598 (Aromatic C=C bend), 1575 (Aromatic C=C bend), 1456 (Aromatic C=C bend), 1314, 1298, 1271, 736.

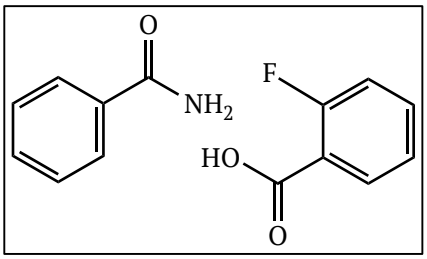
4.8.7 Cocrystals of Benzamide [13]

Benzamide 2-nitrobenzoic acid cocrystal [13.121]



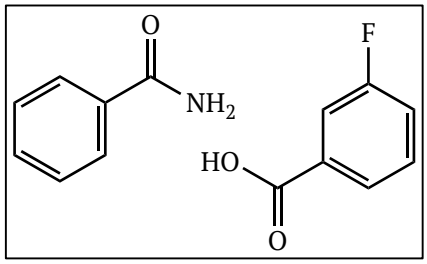
2-Nitrobenzoic acid [**121**] (0.0835 g, 0.50 mmol) and benzamide [**13**] (0.0302 g, 0.25 mmol) were used. Colourless plate crystals of **13.121** were obtained from ethanol. DSC (endotherm); 87.5-90 °C; ν_{max} (ATR)/cm^{-1}; 3450, 3365 (w) (N-H stretches), 1698 (C=O), 1659 (C=O), 1529 (Aromatic C=C bend), 1371 (N=O), 1290; Crystal data for **13.121**: $C_{21}H_{17}O_9N_3$, *Mr* = 455.38, triclinic, *P*-1, a = 7.988(3) Å, b = 11.004(5) Å , c = 12.725(6) Å, α = 73.939(10)°, β = 75.605(1)°, γ = 89.042(11)°, *V* = 1039.5(8) Å^3, *Z* = 2, D_c = 1.455 g cm^{-3}, F_{000} = 472, Mo Kα radiation, λ = 0.71073 Å, *T* = 300(2) K, $2\theta_{max}$ = 26.59°, μ = 0.116 mm^{-1}, 17332 reflections collected, 4182 Unique (R_{int} = 0.804), final GooF = 0.990, R_1 = 0.0654, [1843 obs. Data: $I>2\sigma(I)$]; wR_2 = 0.2286 (all data).

Benzamide 2-fluorobenzoic acid cocrystal [13.84]



2-Fluorobenzoic acid [**84**] (0.0350 g, 0.25 mmol) and benzamide [**13**] (0.0302 g, 0.25 mmol) were used. DSC (endotherm) 69-73 °C; ν_{max} (ATR)/cm^{-1}; 3419, 3366 (w) (N-H stretches), 3171 (w) (C-H stretch), 1694 (C=O), 1659 (C=O), 1466 (Aromatic C=C bend), 1303, 1181 (C-O), 1165 (C-O).

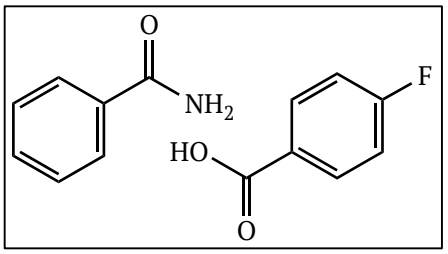
Benzamide 3-fluorobenzoic acid cocrystal [13.85]



3-Fluorobenzoic acid [**85**] (0.0350 g, 0.25 mmol) and benzamide [**13**] (0.0302 g, 0.25 mmol) were used. Colourless block crystals of **13.85** were obtained from ethyl acetate; DSC (endotherm): 80-82 °C; ν_{max} (ATR)/cm^{-1}; 3426, 3334 (w) (N-H stretches), 3210 (w) (C-H stretches), 1682 (C=O), 1632 (C=O),

1557 (Aromatic C=C bend), 1449 (Aromatic C=C bend), 1282 (C-F), 1002, 885; Crystal data for **13.85**: $C_{14}H_{12}O_3NF$, *Mr* = 261.25, triclinic, *P*-1, a = 5.214(4) Å, b = 8.802(7) Å , c = 14.559(12) Å, α = 101.620(19)°, β = 94.456(18)°, γ = 94.83(2)°, *V* = 649.1(9) Å^3, *Z* = 2, D_c = 1.337 g cm^{-3}, F_{000} = 272, Mo Kα radiation, λ = 0.71073 Å, *T* = 300(2) K, $2\theta_{max}$ = 26.31°, μ = 0.104 mm^{-1}, 8211 reflections collected, 2523 Unique (R_{int} = 0.0754), final GooF = 0.968, R_1 = 0.0772, [1005 obs. Data: $I>2\sigma(I)$]; wR_2 = 0.2188 (all data).

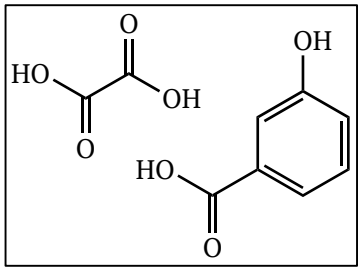
Benzamide 4-fluorobenzoic acid cocrystal [13.86]



4-Fluorobenzoic acid [**86**] (0.0350 g, 0.25 mmol) and benzamide [**13**] (0.0302 g, 0.25 mmol) were used. DSC (endotherm) 106-108 °C (minor), 110-111 °C (major); mp 112-114 °C; ν_{max} (ATR)/cm^{-1}; 3460, 3328 (w) (N-H stretches), 3218 (C-H stretch), 1681 (C=O), 1629 (C=O), 1592, 1310 (C-F), 1181, 1234 (C-O), 774.

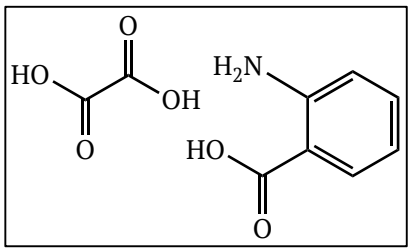
4.8.8 Cocrystals of Oxalic Acid [110]

Oxalic acid 3-hydroxybenzoic acid cocrystal [110.82]



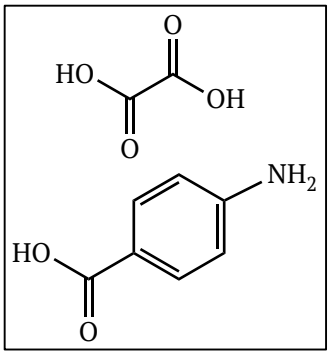
3-Hydroxybenzoic acid [**82**] (0.0345 g, 0.25 mmol) and oxalic acid dihydrate [**110**] (0.0315 g, 0.25 mmol) were used. DSC (endotherm) 99-101 °C (residual oxalic acid), 161-163 °C; ν_{max} (ATR)/cm^{-1}; 3489, 3416, 3341 (w) (O-H stretches), 1681 (C=O), 1598 (Aromatic C=C bend), 1461 (Aromatic C=C bend), 1426 (Aromatic C=C bend), 1233, 1162 (C-O), 920, 723; The observation of residual oxalic acid in DSC suggests a 2:1 ratio of 3-hydroxybenzoic acid to oxalic acid in the cocrystal, time constraints did not allow for further investigation into this cocrystal.

Oxalic acid 2-aminobenzoic acid cocrystal [110.77]



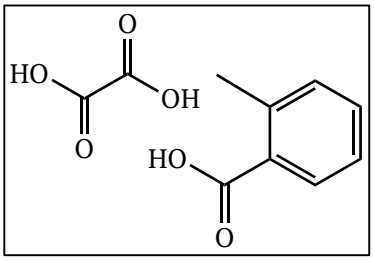
2-Aminobenzoic acid [**77**] (0.0343 g, 0.25 mmol) and oxalic acid dihydrate [**117**] (0.0315 g, 0.25 mmol) were used. mp 149-151 °C; ν_{max} (ATR)/cm^{-1}; 1698 (C=O), 1562 (Aromatic C=C bend), 1486 (Aromatic C=C bend), 1287, 1255, 745.

Oxalic acid 4-aminobenzoic acid cocrystal [110.2]



4-Aminobenzoic acid [**2**] (0.0343 g, 0.25 mmol) and oxalic acid [**117**] (0.0315 g, 0.25 mmol) were used. mp 160 °C decomp.; ν_{max} (ATR)/cm^{-1}; 3335 (w) (O-H stretch),1712 (C=O), 1686 (C=O), 1606, 1233 (C-O), 1171 (C-O).

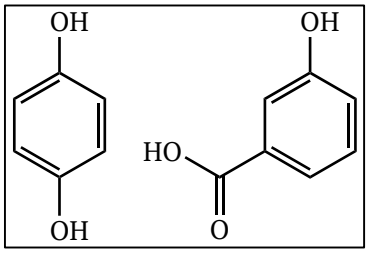
Oxalic acid 2-methylbenzoic acid cocrystal [110.79]



2-Methylbenzoic Acid [**79**] (0.0341 g, 0.25 mmol) and oxalic acid [**117**] (0.0315 g, 0.25 mmol) were used. mp 96-99 °C; ν_{max} (ATR)/cm^{-1}; 3418 (w) (O-H stretch), 1681 (C=O), 1270, 1245 (C-O), 916 (w), 690.

4.8.9 Cocrystals of Hydroquinone [5]

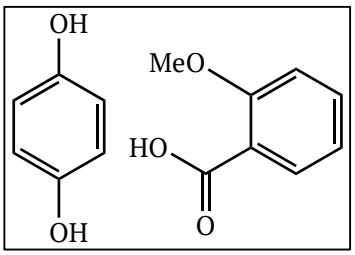
Hydroquinone 3-hydroxybenzoic acid cocrystal [5.82]



3-Hydroxybenzoic acid [**82**] (0.0345 g, 0.25 mmol) and hydroquinone [**5**] (0.0275 g, 0.25 mmol) were used. DSC (endotherm): 152-155 °C; ν_{max} (ATR)/cm^{-1}; 3246 (w) (O-H Stretch), 1684 (C=O),

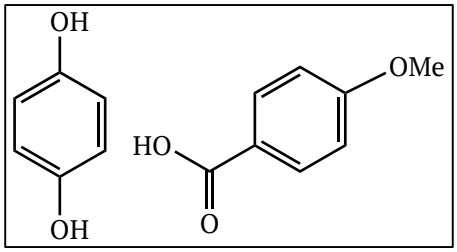
1599 (Aromatic C=C bend), 1461 (Aromatic C=C bend), 1234 (C-O), 1207 (C-O), 1190 (C-O), 920, 756.

Hydroquinone 2-methoxybenzoic acid cocrystal [5.123]



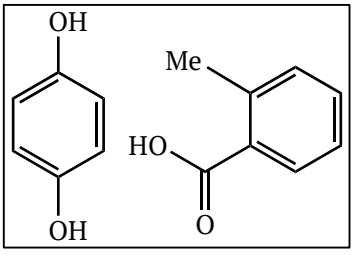
2-Methoxybenzoic acid [**123**] (0.0382 g, 0.25 mmol) and hydroquinone [**5**] (0.0275 g, 0.25 mmol) were used. mp 99-100 °C; ν_{max} (ATR)/cm^{-1}; 1691 (C=O), 1667 (C=O), 1463, 1252, 1207 (C-O), 1191 (C-O), 1154 (C-O), 1087 (C-O), 1049, 826, 757.

Hydroquinone 4-methoxybenzoic acid cocrystal [5.124]



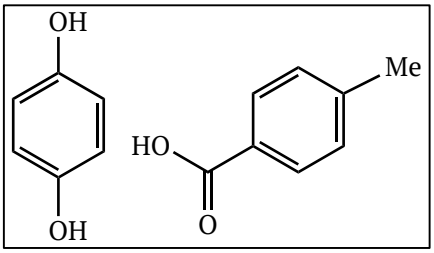
4-Methoxybenzoic acid [**124**] (0.0382 g, 0.25 mmol) and hydroquinone [**5**] (0.0275 g, 0.25 mmol) were used. mp 97-99 °C; ν_{max} (ATR)/cm^{-1}; 3284 (w) (O-H Stretch), 1681 (C=O). 1661 (C=O), 1652 (C=O), 1491 (Aromatic C=C bend), 1462 (Aromatic C=C bend), 1269, 1251, 1167 (C-O), 1103 (C-O), 1018, 759.

Hydroquinone 2-methylbenzoic acid cocrystal [5.79]



2-Methylbenzoic Acid [**79**] (0.0341 g, 0.25 mmol) and Hydroquinone [**5**] (0.0275 g, 0.25 mmol) were used. DSC (endotherm) 89°C (minor), 101-103 °C (major); mp 97-100 °C; ν_{max} (ATR)/cm^{-1}; 1681 (C=O), 1490 (Aromatic C=C bend), 1299, 1119 (C-O), 917, 806, 736.

Hydroquinone 4-methylbenzoic acid cocrystal [5.81]



4-Methylbenzoic Acid [**81**] (0.0341 g, 0.25 mmol) and Hydroquinone [**5**] (0.0275 g, 0.25 mmol) were used. DSC (endotherm) 99-101 °C; ν_{max} (ATR)/cm^{-1}; 3489 3420 (w) (O-H Stretch), 1669 (C=O), 1611, 1417 (Aromatic C=C bend), 1283, 1245 (C-O), 1183 (C-O), 753.

Chapter 4

References

4.9 References

(1) Wicker, J. G. P.; Cooper, R. I. *CrystEngComm* **2015**, *17*, 1927–1934.

(2) Samuel, A. L. *IBM J. Res. Dev.* **1959**, *3*, 210–229.

(3) Sathya, R.; Abraham, A. *Int. J. Adv. Res. Artif. Intell.* **2013**, *2*, 34–38.

(4) Friedman, J. H. *J. Classif.* **2006**, *23*, 175–197.

(5) http://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/. Machine Learning.

(6) Cortes, C.; Vapnik, V. *Mach. Learn.* **1995**, *20*, 273–297.

(7) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, É. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

(8) Landrum, G. RDKit, Open-Source Cheminformatics. http://www.rdkit.org.

(9) R., G. C.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. *Acta Crystallogr. Sect. B* **2016**, *72*, 171–179.

(10) Wood, P. A.; Feeder, N.; Furlow, M.; Galek, P. T. A.; Groom, C. R.; Pidcock, E. *CrystEngComm* **2014**, *16*, 5839–5848.

(11) Reilly, A. M.; Cooper, R. I.; Adjiman, C. S.; Bhattacharya, S.; Boese, A. D.; Brandenburg, J. G.; Bygrave, P. J.; Bylsma, R.; Campbell, J. E.; Car, R.; Case, D. H.; Chadha, R.; Cole, J. C.; Cosburn, K.; Cuppen, H. M.; Curtis, F.; Day, G. M.; DiStasio, R. A.; Dzyabchenko, A.; Van Eijck, B. P.; Elking, D. M.; Van Den Ende, J. A.; Facelli, J. C.; Ferraro, M. B.; Fusti-Molnar, L.; Gatsiou, C. A.; Gee, T. S.; De Gelder, R.; Ghiringhelli, L. M.; Goto, H.; Grimme, S.; Guo, R.; Hofmann, D. W. M.; Hoja, J.; Hylton, R. K.; Iuzzolino, L.; Jankiewicz, W.; De Jong, D. T.; Kendrick, J.; De Klerk, N. J. J.; Ko, H. Y.; Kuleshova, L. N.; Li, X.; Lohani, S.; Leusen, F. J. J.; Lund, A. M.; Lv, J.; Ma, Y.; Marom, N.; Masunov, A. E.; McCabe, P.; McMahon, D. P.; Meekes, H.; Metz, M. P.; Misquitta, A. J.; Mohamed, S.; Monserrat, B.; Needs, R. J.; Neumann, M. A.; Nyman, J.; Obata, S.; Oberhofer, H.; Oganov, A. R.; Orendt, A. M.; Pagola, G. I.; Pantelides, C. C.; Pickard, C. J.; Podeszwa, R.; Price, L. S.; Price, S. L.; Pulido, A.; Read, M. G.; Reuter, K.; Schneider, E.; Schober, C.; Shields, G. P.; Singh, P.; Sugden, I. J.; Szalewicz, K.; Taylor, C. R.; Tkatchenko, A.; Tuckerman, M. E.; Vacarro, F.;

Vasileiadis, M.; Vazquez-Mayagoitia, A.; Vogt, L.; Wang, Y.; Watson, R. E.; De Wijs, G. A.; Yang, J.; Zhu, Q.; Groom, C. R. *Acta Crystallogr. Sect. B* **2016**, *72*, 439–459.

(12) Karamertzanis, P. G.; Kazantsev, A. V; Issa, N.; Welch, G. W. A.; Adjiman, C. S.; Pantelides, C. C.; Price, S. L.; Gareth, W. A.; Adjiman, C. S.; Pantelides, C. C.; Price, S. L. *Cryst. Growth Des.* **2009**, *9*, 442–453.

(13) Fábián, L. *Cryst. Growth Des.* **2009**, *9*, 1436–1443.

(14) Grecu, T.; Hunter, C. A.; Gardiner, E. J.; McCabe, J. F. *Cryst. Growth Des.* **2014**, *14*, 165–171.

(15) Hunter, C. A. *Angew. Chemie - Int. Ed.* **2004**, *43*, 5310–5324.

(16) Abramov, Y. A.; Loschen, C.; Klamt, A. *J. Pharm. Sci.* **2012**, *101*, 3687–3697.

(17) Delori, A.; Galek, P. T. a.; Pidcock, E.; Patni, M.; Jones, W. *Cryst. Eng. Comm.* **2013**, *15*, 2916.

(18) Galek, P. T.; Fábián, L.; Motherwell, W. D. S.; Allen, F. H.; Feeder, N. *Acta Crystallogr. B.* **2007**, *63*, 768–782.

(19) Macrae, C. F.; Bruno, I. J.; Chisholm, J. A.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; van de Streek, J.; Wood, P. A. J. *J. Appl. Crystallogr.* **2008**, *41*, 466–470.

(20) Delori, A.; Friscic, T.; Jones, W.; Friščić, T.; Jones, W. *CrystEngComm* **2012**, *14*, 2350–2362.

(21) Dai, Y.-M.; Huang, J.-F.; Shen, H.-Y. *Acta Crystallogr. Sect. E Struct. Reports Online* **2005**, *61*, o3919–o3920.

(22) Aakeröy, C. B.; Desper, J.; Helfrich, B. A. *CrystEngComm* **2004**, *6*, 19–24.

(23) Zhu, Z.; Wang, F.-Q.; Zhao, Y.-N. *Acta Crystallogr. Sect. E Struct. Reports Online* **2010**, *66*, o546.

(24) Seaton, C. C.; Parkin, A. *Cryst. Growth Des.* **2011**, *11*, 1502–1511.

(25) Beatty, A. M.; Helfrich, B. A.; Nieuwenhuyzen, M. *Cryst. Growth Des.* **2003**, *3*, 159–165.

(26) Bowers, J. R.; Hopkins, G. W.; Yap, G. P. A.; Wheeler, K. A. *Cryst. Growth Des.* **2005**, *5*, 727–736.

(27) Tan, T.-F.; Han, J.; Pang, M.-L.; Song, H.-B.; Ma, Y.-X.; Meng, J.-B. *CSD Entry DAQZIF01* **2006**, *6*, 1186.

(28) Blagden, N.; Berry, D. J.; Parkin, A.; Javed, H.; Ibrahim, A.; Gavan, P. T.; De

Matos, L. L.; Seaton, C. C. *New J. Chem.* **2008**, *32*, 1659.

(29) Moragues-Bartolome, A. M.; Jones, W.; Cruz-Cabeza, A. J. *CrystEngComm* **2012**, *14*, 2552–2559.

(30) Berry, D. J.; Seaton, C. C.; Clegg, W.; Harrington, R. W.; Coles, S. J.; Horton, P. N.; Hursthouse, M. B.; Storey, R.; Jones, W.; Friščić, T.; Blagden, N. *Cryst. Growth Des.* **2008**, *8*, 1697–1712.

(31) Hathwar, V. R.; Pal, R.; Guru Row, T. N. *Cryst. Growth Des.* **2010**, *10*, 3306–3310.

(32) Berry, D.; Blagden, N.; Coles, S. J.; Hursthouse, M. B. *Univ. Southampton, Cryst. Struct. Rep. Arch.* **2007**, 295.

(33) McMahon, J. A.; Bis, J. A.; Vishweshwar, P.; Shattock, T. R.; McLaughlin, O. L.; Zaworotko, M. J. *Z. fur Krist.* **2005**, *220*, 340–350.

(34) Zhou, Z.; Chan, H. M.; Sung, H. H. Y.; Tong, H. H. Y.; Zheng, Y. *Pharm. Res.* **2016**, *33*, 1030–1039.

(35) Tothadi, S. *CrystEngComm* **2014**, *16*, 7587–7597.

(36) Gellert, R. W.; Hsu, I.-N. *Acta Crystallogr. Sect. A Cryst. Phys. Diffr. Theor. Crystallogr.* **1981**, *37*, C93.

(37) Emsley, J.; Reza, N. M. *J. Crystallogr. Spectrosc. Res.* **1986**, *16*, 57–69.

(38) Hsu, I. N.; Gellert, R. W. *J. Crystallogr. Spectrosc. Res.* **1983**, *13*, 43–48.

(39) Elbagerma, M. A.; Edwards, H. G. M.; Munshi, T.; Scowen, I. J. *Anal. Bioanal. Chem.* **2010**, *397*, 137–146.

(40) Aakeroy, C. B.; Beatty, A. M.; Helfrich, B. A. *J. Am. Chem. Soc.* **2002**, *124*, 14425–14432.

(41) Shattock, T. R.; Arora, K. K.; Vishweshwar, P.; Zaworotko, M. J. *Cryst. Growth Des.* **2008**, *8*, 4533–4545.

(42) Lapidus, S. H.; Stephens, P. W.; Arora, K. K.; Shattock, T. R.; Zaworotko, M. J. *Cryst. Growth Des.* **2010**, *10*, 4630–4637.

(43) Arenas-García, J. I.; Herrera-Ruiz, D.; Mondragón-Vásquez, K.; Morales-Rojas, H.; Höpfl, H. *Cryst. Growth Des.* **2010**, *10*, 3732–3742.

(44) Lemmerer, A.; Adsmond, D. A.; Bernstein, J. *Cryst. Growth Des.* **2011**, *11*, 2011–2019.

(45) Vishweshwar, P.; Nangia, A.; Lynch, V. M. *CrystEngComm* **2003**, *5*, 164.

(46) Mukherjee, A.; Desiraju, G. R. *Chem. Commun.* **2011**, *47*, 4090–4092.

(47) Hathwar, V. R.; Thakur, T. S.; Dubey, R.; Pavan, M. S.; Guru Row, T. N.; Desiraju, G. R. *J. Phys. Chem. A* **2011**, *115*, 12852–12863.

(48) Arman, H. D.; Tiekink, E. R. T. *Acta Crystallogr. Sect. E Struct. Reports Online* **2013**, *69*, o1447.

(49) Fischer, F.; Joester, M.; Rademann, K.; Emmerling, F. *Chem. - A Eur. J.* **2015**, *21*, 14969–14974.

(50) Lhengwan, P.; Achiwawanich, S.; Duangthongyou, T. *Acta Crystallogr. Sect. E Struct. Reports Online* **2012**, *68*, o2569.

(51) Lynch, D. E.; Chatwin, S.; Parsons, S. *Cryst. Eng.* **1999**, *2*, 137–144.

(52) Bendjeddou, L.; Farah, S.; Cherouana, A. *Acta Crystallogr. Sect. E Struct. Reports Online* **2009**, *65*, o1839–o1840.

(53) Lou, B.; Hu, S. *J. Chem. Crystallogr.* **2011**, *41*, 1663–1668.

(54) Mukherjee, A.; Desiraju, G. R. *Cryst. Growth Des.* **2014**, *14*, 1375–1385.

(55) Wang, R.; Jiang, F.; Zhou, Y.; Han, L.; Hong, M. *Inorganica Chim. Acta* **2005**, *358*, 545–554.

(56) Smith, G.; Baldry, B. K. E.; C, A. K. A. B.; C, C. H. L. K. *Aust. J. Chem.* **1997**, *50*, 727–736.

(57) Qian, X.-Y.; Liu, F.-Z. *Acta Crystallogr. Sect. E Struct. Reports Online* **2012**, *68*, o597.

(58) Ramon, G.; Davies, K.; Nassimbeni, L. R. *CrystEngComm* **2014**, *16*, 5802–5810.

(59) Sudhakar, P.; Srivijaya, R.; Sreekanth, B. R.; Jayanthi, P. K.; Vishweshwar, P.; Babu, M. J.; Vyas, K.; Iqbal, J. *J. Mol. Struct.* **2008**, *885*, 45–49.

(60) Ganduri, R.; Cherukuvada, S.; Guru Row, T. N. *Cryst. Growth Des.* **2015**, *15*, 3474–3480.

(61) Jebas, S. R.; Balasubramanian, T. *Acta Crystallogr. Sect. E Struct. Reports Online* **2006**, *62*, o5621–o5622.

(62) Luo, Y. H.; Xu, B.; Sun, B. W. *J. Cryst. Growth* **2013**, *374*, 88–98.

(63) Tothadi, S.; Desiraju, G. R. *Cryst. Growth Des.* **2012**, *12*, 6188–6199.

(64) Edwards, M. R.; Jones, W.; Motherwell, W. D. S. *Cryst. Eng. Comm.* **2006**, *8*, 545.

(65) Stokes, S. P.; Lawrence, S. E. *Unpubl. Work*.

(66) Perlovich, G. L. *CrystEngComm* **2015**, *17*, 7019–7028.

(67) Kuhn, R.; Vetter, H. *Berichte der Dtsch. Chem. Gesellschaft [Abteilung] B Abhandlungen* **1935**, *68B*, 2374–2375.

(68) Kursanov, D. N.; Kichkina, A. S. *Zhurnal Obs. Khimii* **1935**, *5*, 1342–1347.

(69) Bakunin, M.; Angrisani, T. *Gazz. Chim. Ital.* **1915**, *45*, 197–204.

(70) Holleman, A. F. *Recl. des Trav. Chim. des Pays-Bas la Belgique* **1907**, *25*, 330–333.

(71) Senier, A.; Forster, R. B. *J. Chem. Soc. Trans.* **1915**, *107*, 1168–1173.

(72) Jones, T. G. H. *Proc. Roy. Soc. Queensland.* **1934**, *45*, 38–40.

(73) Auwers, K.; Roth, W. A. *Justus Liebigs Ann. Chem.* **1910**, *373*, 239–248.

(74) Wallach, O. *Justus Liebigs Ann. Chem.* **1908**, *360*, 26–81.

(75) Gastaldi, C.; Cherchi, F. *Gazz. Chim. Ital.* **1915**, *45*, 251–275.

(76) Tsuda, K.; Hayatsu, R. *Yakugaku Zasshi* **1952**, *72*, 1303–1310.

(77) Inoue, Y.; Onodera, K.; Kitaoka, S. *Nippon Nogei Kagaku Kaishi* **1952**, *25*, 291–293.

(78) Wibaut, J. P.; Dingemanse, E.; Wilbaut, J. P.; Dingemanse, E. *Recl. des Trav. Chim. des Pays-Bas la Belgique* **1923**, *42*, 240–250.

(79) Zellner, J.; Zikmunda, E. *Monatshefte fuer Chemie* **1930**, *56*, 200–203.

(80) Lepsius, R. *Justus Liebigs Ann. Chem.* **1914**, *406*, 11–21.

(81) Smith, H. L. *Analyst* **1916**, *41*, 3–6.

(82) Semishin, V. I. *Zhurnal Obs. Khimii* **1939**, *9*, 83–85.

(83) Robertson, J. M.; Ubbelohde, A. R. *Nature* **1937**, *139*, 504–505.

(84) Lee Lewis, W.; Cheetham, H. C. *J. Am. Chem. Soc.* **1921**, *43*, 2117–2121.

(85) Angeli, A. *Atti della Accad. Naz. dei Lincei, Cl. di Sci. Fis. Mat. e Nat. Rend.* **1914**, *23*, 557–570.

(86) Biilmann, E.; Lund, H. *Ann. di Chim. Appl.* **1922**, *18*, 263–282.

(87) Rolfe, H. M. *J. Cosmet. Dermatol.* **2014**, *13*, 324–328.

(88) Chen, A. C.; Damian, D. L. *Australas. J. Dermatol.* **2014**, *55*, 169–175.

(89) Wohlrab, J.; Kreft, D. *Skin Pharmacol. Physiol.* **2014**, *27*, 311–315.

(90) Brittain, H. G. *Profiles Drug Subst. Excipients Relat. Methodol.* **2011**, *36*, 361–381.

(91) Brittain, H. G. *Cryst. Growth Des.* **2012**, *12*, 5823–5832.

(92) Hino, T.; Ford, J. L.; Powell, M. W. *Thermochim. Acta* **2001**, *374*, 85–92.

(93) Miwa, Y.; Mizuno, T.; Tsuchida, K.; Taga, T.; Iwata, Y. *Acta Crystallogr. Sect.*

B Struct. Sci. **1999**, *55*, 78–84.

(94) Li, J.; Bourne, S.; Caira, M. R. *Chem. Commun.* **2011**, *47*, 1530–1532.

(95) Aakeröy, C. B.; Salmon, D. J. *CrystEngComm* **2005**, *7*, 439–448.

(96) Eccles, K. S.; Elcoate, C. J.; Maguire, A. R.; Lawrence, S. E. *Cryst. Growth Des.* **2011**, *11*, 4433–4439.

(97) Das, B.; Baruah, J. B. *Cryst. Growth Des.* **2011**, *11*, 5522–5532.

(98) Tothadi, S.; Desiraju, G. R. *Philos. Trans. A. Math. Phys. Eng. Sci.* **2012**, *370*, 2900–2915.

(99) Eccles, K. S.; Deasy, R. E.; Fabian, L.; Braun, D. E.; Maguire, A. R.; Lawrence, S. E. *CrystEngComm* **2011**, *13*, 6923–6925.

(100) Ohlan, R.; Narasimhan, B.; Ohlan, S.; Narang, R.; Judge, V. *Org. Commun.* **2008**, *1*, 24–32.

(101) Przyblek, M.; Ziolkowska, D.; Kobierski, M.; Mroczynska, K.; Cysewski, P. *J. Cryst. Growth* **2016**, *433*, 128–138.

(102) Chaudhari, S. R.; Suryaprakash, N. *J. Mol. Struct.* **2012**, *1016*, 163–168.

(103) Jankowski, W.; Gdaniec, M.; Połoński, T. *Acta Crystallogr. Sect. C Cryst. Struct. Commun.* **2006**, *62*, 492–494.

Chapter 5

Future Work

5.1 Future Work

5.1.1 Synthesis and Solid State Characterisation of Aryl Primary Sulfinamides – Future Work

Chapter 2 has described the crystalline landscape of aryl primary sulfinamides in depth. The knowledge gained in this chapter creates opportunity for further exploration of this interesting functional group. The first area that could be investigated as an extension of this work would be to synthesis and characterise a related series of alkyl primary sulfinamides, each with a range of competing donor and acceptor moieties on the alkyl chain.

By varying chain length, and the presence of competitive hydrogen bond donors and/or acceptors, one could gain further insight into the robustness of the strong N-H···O=S hydrogen bonds, which have been identified as significant structure defining features for the aryl series of compounds. Building upon this, the relative stability of this related series of alkyl primary sulfinamides could be investigated from solution crystallization. It would be interesting to observe how the greater conformational flexibility associated with an alkyl chain would affect the S=O group with respect to hydrolysis.

In keeping with the long term goal of the project, it would be interesting to develop our understanding of aryl primary sulfinamides as cocrystallization targets. There are two potential directions in which to investigate primary sulfinamides for cocrystallization. The first route for this would be to build upon the previous work within our research group, which has identified that halogen bond donors such as 1,4-diiodotetrafluorobenzene [**132**] are suitable coformers for use with the primary sulfinamide moiety.[1] This knowledge could be extended to determine the scope of similar coformers as halogen bond donors for a series of primary sulfinamides (be they alkyl, aryl or both). Since cocrystallization can have a positive effect on stability,[2] it would be interesting to observe the relative rates of hydrolysis of these materials if a cocrystal was successfully prepared.

A second route for this cocrystallization project would be to investigate the potential for cocrystallization of those sulfinamides which were observed to be

more stable, with a series of carboxylic acid coformers. Analysis of Hunter's values[3] would suggest that carboxylic acids should be a suitable hydrogen bond donor for the primary sulfinamide functional group. However, given the hydrolytic senstivity of many of these materials, it would be wise to develop cocrystallization screening methods that are not solution-based, such as neat grinding, sublimation or melt cocrystallization.

5.1.2 Cocrystallization of Salsalate – Future Work

Salsalate [**58**] has proven to be a very interesting, and surprising API target for cocrystallization studies. Building upon the suite of cocrystals developed in chapter 3, it would be of value to target salsalate with a range of pyridine-containing API molecules in the development of dual-drug cocrystals.

A more interesting aspect of the solid state characteristics of salsalate was that of the reactive cocrystallization phenomena observed with 4-methylbenzamide [**66**] and 2-fluorobenzamide [**70**]. The instability of **58** in the presence of these materials could have significant impact upon drug formulation of salsalate, since compounds **66** and **70** are both GRAS coformers, suitable for use in drug formulations. The reactive cocrystallization would indicate that the stability of the **58** in the presence of these materials may not be sufficient as to allow for effective transition of **58** through the GI tract before forming salicylate [**59**], thereby increasing the likelihood of gastric injury in such a formulation.

Further research is required in this instance, firstly to develop an understanding of the exact mechanism of this reactive cocrystallization, and secondly, to probe the potential utility of such a phenomenon; work on this is ongoing within our research group.

5.1.3 Machine learning methods for cocrystallization prediction – Future Work

The machine learning algorithm developed in chapter 4 has significant applications in a real-world setting. The future for this research would see the algorithm disseminated amongst the academic community for use in cocrystallization screening. Initially, it would be beneficial to test the utility of the algorithm in directing a cocrystallization screen for a target API. The ability of the algorithm to correctly rank coformers in order of likelihood to cocrystallize presents significant advantages for the quick and successful preparation of cocrystals.

It is hoped that the algorithm will be used effectively going forward to direct cocrystallization screens, particularly since the coformers used for preparation of the data matrix are now a standard set of acid/amide coformers used within our research group for cocrystallization screening.

Finally, additional work can be completed by analysing the decision trees for the algorithm to gain insight into the determinant factors that the algorithm uses for prediction of cocrystals. This could present some interesting results which could be built upon for fine-tuned coformer selection in the future.

Chapter 5

References

5.2 References

(1) Eccles, K. S.; Morrison, R. E.; Daly, C. A.; O'Mahony, G. E.; Maguire, A. R.; Lawrence, S. E. *CrystEngComm* **2013**, *15*, 7571–7575.

(2) Wang, Z. Z.; Chen, J. M.; Lu, T. B. *Cryst. Growth Des.* **2012**, *12*, 4562–4566.

(3) Hunter, C. A. *Angew. Chemie - Int. Ed.* **2004**, *43*, 5310–5324.